Capacity Results for Binary Fading Interference Channels with Delayed CSIT

Alireza Vahid, Mohammad Ali Maddah-Ali, and A. Salman Avestimehr

Abstract

To study the effect of lack of up-to-date channel state information at the transmitters (CSIT), we consider two-user binary fading interference channels with Delayed-CSIT. We characterize the capacity region for such channels under homogeneous assumption where channel gains have identical and independent distributions across time and space, eliminating the possibility of exploiting time/space correlation. We introduce and discuss several novel coding opportunities created by outdated CSIT, which can enlarge the achievable rate region. The capacity-achieving scheme relies on accurate combination, concatenation, and merging of these opportunities, depending on the channel statistics. The outer-bounds are based on an extremal inequality we develop for a binary broadcast channel with Delayed-CSIT. We further extend the results and characterize the capacity region in the case that output feedback links from the receivers to the transmitters are available in addition to the delayed knowledge of the channel state information.

I. Introduction

The history of studying the impact of feedback channel in communication systems traces back to Shannon [3], and ever since, there have been extensive efforts to discover new techniques that exploit feedback channels in order to benefit wireless networks. In today's wireless networks, one of the main objectives in utilizing feedback channels is to provide the transmitters with the knowledge of the channel state information (CSI). In slow-fading networks, this task could have been carried on with negligible overhead. However, as wireless networks started growing in size, as mobility became an inseparable part of networks, and as fast-fading networks started playing a more important role, the availability of up-to-date channel state information at the transmitters (CSIT) has become a challenging task to accomplish. Specifically, in fast-fading scenarios, the coherence time of the channel is smaller than the delay of the feedback channel, and as a result, providing the transmitters with up-to-date channel state information is practically infeasible.

As a result, there has been a recent growing interest in studying the effect of lack of up-to-date channel state information at the transmitters in wireless networks. In particular, in the context of multiple-input single-output (MISO) broadcast channels (BC), it was recently shown that even completely stale CSIT (a.k.a. Delayed-CSIT) can still be very useful and can change the scale of the capacity, measured by the degrees of freedom (DoF) [4]. A key idea behind the scheme proposed in [4] is that instead of predicting future channel state information, transmitters should focus on the side-information provided in the past signaling stages via the feedback channel, and try to create signals that are of *common interest* of multiple receivers. Hence, we can increase spectral efficiency by retransmission of such signals of common interest.

There have also been several recent works in the literature to understand the impact of Delayed-CSIT in wireless networks with distributed transmitters. This includes studying the DoF region of multi-antenna two-user Gaussian IC and X channel [5], [6], k-user Gaussian IC and X channel [7], [8], and multi-antenna two-user Gaussian IC with Delayed-CSIT and Shannon feedback [9], [10]. In particular, the DoF region of multi-antenna two-user Gaussian IC has been characterized in [11], and it has been shown that the k-user Gaussian IC and X channels can still achieve more than one DoF with Delayed-CSIT [7], [8] (for $k > 2$).

A. Vahid and A. S. Avestimehr are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA. Email: av292@cornell.edu and avestimehr@ece.cornell.edu.

Mohammad Ali Maddah-Ali is with Bell Labs, Alcatel-Lucent, Holmdel, NJ, USA. Email: mohammadali.maddah-ali@alcatel-lucent.com.

The work of A. S. Avestimehr and A. Vahid is in part supported by NSF Grants CAREER-0953117, CCF-1161720, NETS-1161904, AFOSR Young Investigator Program Award, and ONR award N000141310094.

Preliminary parts of this work were presented at the 2011 Allerton Conference on Communication, Control, and Computing [1], and the 2012 International Symposium on Information Theory (ISIT) [2].

A major challenge that arises in interference channels with Delayed-CSIT is that in such networks the transmitter has no longer access to all the transmit signals in the network. In fact, each transmitter has *only* access to its own interference contribution. Therefore, unlike BC in which the task of creating signals of common interest could be simply done at a single transmitter that has access to *all* messages, exploiting Delayed-CSIT becomes much more challenging. In order to shed light on fundamental limits of communications with Delayed-CSIT in interference channels, in this paper we focus on a binary fading model as described below.

We consider a two-user interference channel as illustrated in Figure 1. In this network, the channel gains at each time instant are either 0 or 1 according to some Bernoulli distribution, and are independent from each other and *over time*. The input and output signals are also in the binary field and if two signals arrive simultaneously at a receiver, then the receiver obtains the XOR of them. We shall refer to this network as the two-user Binary Fading Interference Channel (BFIC).

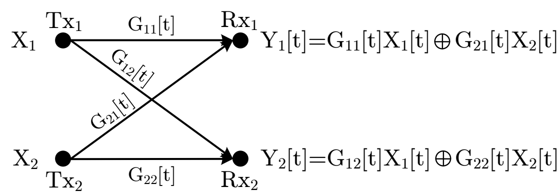


Fig. 1. Binary fading channel model for a two-user interference channel. The channel gains, the transmitted signals and the received signals are in the binary field. The channel gains are distributed as i.i.d. Bernoulli random variables. The channel gains are independent across time so that the transmitters cannot predict future based on the past channel state information.

As the main motivation, we study the two-user BFIC as a stepping stone towards understanding the capacity of more complicated fading interference channels with Delayed-CSIT. Lately, the linear deterministic model introduced in [12], has been utilized to bridge the results from deterministic networks into Gaussian networks (*e.g.*, [12]–[18]). In the linear deterministic model, there is a non-negative integer representing the channel gain from a transmitter to a receiver. Hence, one can view the binary fading model as a fading interpretation of the linear deterministic model where the non-negative integer associated to each link is either 0 or 1. The binary fading model is also motivated by a simple shadow fading environment, in which each link is either "on" or "off" (according to the shadow fading distribution), and the multiple access (MAC) is modeled such that if two signals are transmitted simultaneously, and the links between the corresponding transmitters and the receiver are not in deep fade, then a linear combination of the signals is available to the receiver.

In this work, we fully characterize the capacity region of the two-user BFIC with Delayed-CSIT. We introduce and discuss several novel coding opportunities, created by outdated CSIT, which can enlarge the achievable rate region. In particular, we propose a new transmission strategy, which is carried on over several phases. Each channel realization creates multiple coding opportunities which can be exploited in the next phases, to improve the rate region. However, we observe that *merging* or *concatenating* some of the opportunities can offer even more gain. To achieve the capacity region, we find the most efficient arrangement of combination, concatenation, and merging of the opportunities, depending on the channel statistics. This can take up to five phases of communication for a two-user channel. For converse arguments, we start with a genie-aided interference channel and show that the problem can be reduced to some particular form of broadcast channels with Delayed-CSIT. We establish a new extremal inequality for the underlying BC, which leads to a tight outer-bound for the original interference channel. The established inequality provides an outer-bound on how much the transmitter in a BC can favor one receiver to the other using Delayed-CSIT (in terms of the entropy of the received signal at the two receivers).

We also consider the scenario in which output feedback links are available from the receivers to the transmitters on top of the delayed knowledge of the channel state information. We demonstrate how output feedback can be utilized to further improve the achievable rates in terms of both enlarging the capacity region and improving the achievable sum-rate. In addition, output feedback can help us to simplify the achievability strategy significantly. For converse, again the core idea is to reduce the problem to a broadcast channel with Delayed-CSIT and output feedback, and establishing a new extremal inequality for the resultant broadcast channel. The inequality then helps us to prove a tight outer-bound for the original interference channel.

The rest of the paper is organized as follows. In Section II, we formulate our problem. In Section III, we present our main results and illustrate them through an example. We then provide an overview of our main achievability and converse techniques in Section IV. Sections V-X are dedicated to the proof of our main results. Section XI concludes the paper and describes several interesting future directions.

II. Problem Setting

We consider the two-user Binary Fading Interference Channel (BFIC) as illustrated in Figure 2. The channel gain from transmitter Tx_i to receiver Rx_j at time instant t is denoted by $G_{ij}[t]$, $i,j \in \{1,2\}$. We assume that the channel gains are either 0 or 1 (*i.e.* $G_{ij}[t] \in \{0,1\}$), and they are distributed as independent Bernoulli random variables (independent from each other and *over time*). We consider the homogeneous setting where

$$G_{ij}[t] \overset{d}{\sim} \mathcal{B}(p), \qquad i,j = 1,2, \tag{1}$$

for $0 \leq p \leq 1$, and we define $q = 1-p$.

At each time instant t, the transmit signal at Tx_i is denoted by $X_i[t] \in \{0,1\}$, and the received signal at Rx_i is given by

$$Y_i[t] = G_{ii}[t]X_i[t] \oplus G_{\bar{i}i}[t]X_{\bar{i}}[t], \quad i = 1,2, \tag{2}$$

where the summation is in $\mathbb{F}_2$.

Furthermore, the channel state information (CSI) at time instant t is denoted by the quadruple

$$G[t] = (G_{11}[t], G_{12}[t], G_{21}[t], G_{22}[t]). \tag{3}$$

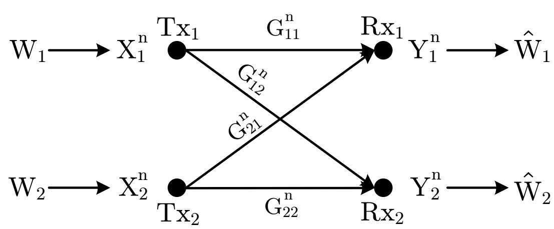


Fig. 2. Two-user Binary Fading Interference Channel (BFIC). The channel gains, the transmit and the received signals are in the binary field. The channel gains are distributed as i.i.d. Bernoulli random variables.

We use the following notations in this paper. We use capital letters to denote random variables (RVs), *e.g.*, $G_{ij}[t]$ is a random variable at time instant t. Furthermore for a natural number k, we set

$$G^k = \left[G[1], G[2], \ldots, G[k]\right]^T. \tag{4}$$

Finally, we set

$$G_{ii}^t X_i^t \oplus G_{\bar{i}i}^t X_{\bar{i}}^t = \left[G_{ii}[1]X_i[1] \oplus G_{\bar{i}i}[1]X_{\bar{i}}[1], \ldots, G_{ii}[t]X_i[t] \oplus G_{\bar{i}i}[t]X_{\bar{i}}[t]\right]^T. \tag{5}$$

In this paper, we consider three models for the available channel state information at the transmitters:

1) Instantaneous-CSIT: In this model, the channel state information G^t is available at each transmitter at time instant t, $t = 1,2,\ldots,n$;
2) No-CSIT: In this model, transmitters only know the distribution from which the channel gains are drawn, but not the actual realizations of them;
3) Delayed-CSIT: In this model, at time instant t, each transmitter has the knowledge of the channel state information up to the previous time instant (*i.e.* G^{t-1}) and the distribution from which the channel gains are drawn, $t = 1,2,\ldots,n$.

We assume that the receivers have instantaneous knowledge of the CSI. Consider the scenario in which Tx_i wishes to reliably communicate message $\mathsf{W}_i \in \{1,2,\ldots,2^{nR_i}\}$ to Rx_i during n uses of the channel, $i = 1,2$. We assume that the messages and the channel gains are *mutually* independent and the messages are chosen uniformly. For each transmitter Tx_i, let message W_i be encoded as X_i^n using the encoding function $f_i(.)$, which depends on

the available CSI at Tx_i. Receiver Rx_i is only interested in decoding W_i, and it will decode the message using the decoding function $\widehat{\mathrm{W}}_i = g_i(Y_i^n, G^n)$. An error occurs when $\widehat{\mathrm{W}}_i \neq \mathrm{W}_i$. The average probability of decoding error is given by

$$\lambda_{i,n} = \mathbb{E}[P[\widehat{\mathrm{W}}_i \neq \mathrm{W}_i]], \quad i = 1, 2, \tag{6}$$

and the expectation is taken with respect to the random choice of the transmitted messages W_1 and W_2. A rate tuple (R_1, R_2) is said to be achievable, if there exists encoding and decoding functions at the transmitters and the receivers respectively, such that the decoding error probabilities $\lambda_{1,n}, \lambda_{2,n}$ go to zero as n goes to infinity.

In addition to the setting described above, we consider a separate scenario in which an output feedback (OFB) link is available from each receiver to its corresponding transmitter[1]. More precisely, we consider a noiseless feedback link of infinite capacity from each receiver to its corresponding transmitter.

Due to the presence of output feedback links, the encoded signal $X_i[t]$ of transmitter Tx_i at time t, would be a function of its own message, previous output sequence at its receiver, and the available CSIT. For instance, with Delayed-CSIT and OFB, we have

$$X_i[t] = f_i[t](\mathrm{W}_i, Y_i^{t-1}, G^{t-1}), \quad i = 1, 2. \tag{7}$$

As stated in the introduction, our goal is to understand the impact of the channel state information and the output feedback, on the capacity region of the two-user Binary Fading Interference Channel. Towards that goal, we consider several scenarios about the availability of the CSIT and the OFB. For all scenarios, we provide exact characterization of the capacity region. In the next section, we present the main results of the paper.

III. Statement of Main Results

In this paper, we basically focus on the following scenarios about the availability of the CSI and the OFB: (1) Delayed-CSIT and no OFB; (2) Delayed-CSIT and OFB; and (3) Instantaneous-CSIT and OFB. In order to illustrate the results, we first establish the capacity region of the two-user BFIC with No-CSIT and Instantaneous-CSIT as our benchmarks.

A. Benchmarks

Our base line is the scenario in which there is no output feedback link from the receivers to the transmitters, and we assume the No-CSIT model. In other words, the only available knowledge at the transmitters is the distribution from which the channel gains are drawn. In this case, it is easy to see that for any input distribution, the two received signals are *statistically* the same, hence

$$\begin{aligned} &I\left(X_1^n; Y_1^n | G^n\right) = I\left(X_1^n; Y_2^n | G^n\right), \\ &I\left(X_2^n; Y_1^n | G^n\right) = I\left(X_2^n; Y_2^n | G^n\right). \end{aligned} \tag{8}$$

Therefore, the capacity region in this case, $\mathcal{C}^{\text{No}-\text{CSIT}}$, is the same as the intersection of the capacity region of the multiple-access channels (MAC) formed at either of the receivers:

$$\mathcal{C}^{\text{No}-\text{CSIT}} = \begin{cases} 0 \leq R_i \leq p, & i = 1, 2, \\ R_1 + R_2 \leq 1 - q^2. \end{cases} \tag{9}$$

The other extreme point on the available CSIT is the Instantaneous-CSIT model. The capacity region in this case is given in the following theorem which is proved in Appendices A and B.

Theorem 1: **[Capacity Region with Instantaneous-CSIT]** The capacity region of the two-user Binary Fading IC with Instantaneous-CSIT (and no output feedback), $\mathcal{C}^{\text{ICSIT}}$, is the set of all rate tuples (R_1, R_2) satisfying

$$\mathcal{C}^{\text{ICSIT}} = \begin{cases} 0 \leq R_i \leq p, & i = 1, 2, \\ R_1 + R_2 \leq 1 - q^2 + pq. \end{cases} \tag{10}$$

Remark 1: Comparing the capacity region of the two-user BFIC with No-CSIT (9) and Instantaneous-CSIT (10), we observe that the bounds on individual rates remain unchanged while the sum-rate outer-bound is increased by

[1]As we will see later, our result holds for the case in which output feedback links are available from each receiver to *both* transmitters.

pq. This increase can be intuitively explained as follows. The outer-bound of $1-q^2$ corresponds to the fraction of time in which at least one of the links to each receiver is equal to 1. Therefore, this outer-bound corresponds to the fraction of time that each receiver gets “useful” signal. This is tight with No-CSIT since each receiver should be able to decode both messages. However, once we move to Instantaneous-CSIT, we can send a private message to one of the receivers by using those time instants in which the link from the corresponding transmitter to that receiver is equal to 1, but that transmitter is not interfering with the other receiver. This corresponds to pq fraction of the time.

Now that we have covered the benchmarks, we are ready to present our main results.

B. Main Results

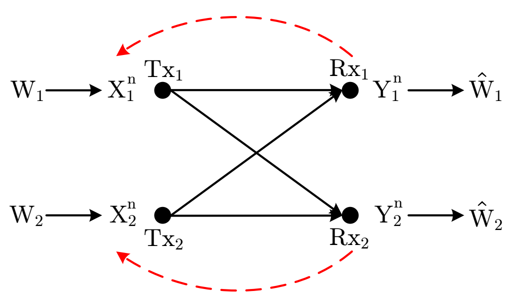
As first step, we consider the Delayed-CSIT model. In this case, the following theorem establishes our result.

Theorem 2: **[Capacity Region with Delayed-CSIT]** The capacity region of the two-user Binary Fading IC with Delayed-CSIT (and no output feedback), $\mathcal{C}^{\mathrm{DCSIT}}$, is the set of all rate tuples (R_1, R_2) satisfying

$$\mathcal{C}^{\mathrm{DCSIT}} = \begin{cases} 0 \leq R_i \leq p, & i=1,2, \\ R_i + (1+q)\, R_{\bar{i}} \leq p\,(1+q)^2\,, & i=1,2. \end{cases} \tag{11}$$

Remark 2: Comparing the capacity region of the two-user BFIC with Delayed-CSIT (11) and Instantaneous-CSIT (10), we can show that for $0 \leq p \leq \left(3-\sqrt{5}\right)/2$, the two regions are equal. However, for $\left(3-\sqrt{5}\right)/2 < p < 1$, the capacity region of the two-user BFIC with Delayed-CSIT is strictly smaller than that of Instantaneous-CSIT. Moreover, we can show that the capacity region of the two-user BFIC with Delayed-CSIT is strictly larger than that of No-CSIT (except for $p=0$ or 1).

Furthermore, since the channel state information is acquired through the feedback channel, it is also important to understand the impact of output feedback on the capacity region of the two-user BFIC with Delayed-CSIT. In the study of feedback in wireless networks, one other direction is to consider the transmitter cooperation created through the output feedback links. In this context, it is well-known that feedback does not increase the capacity of discrete memoryless point-to-point channels [3]. However, feedback can enlarge the capacity region of multi-user networks, even in the most basic case of the two-user memoryless multiple-access channel [19], [20]. In [15], [16], the feedback capacity of the two-user Gaussian IC has been characterized to within a constant number of bits. One consequence of these results is that output feedback can provide an unbounded capacity increase. This is in contrast to point-to-point and multiple-access channels where feedback provides no gain and bounded gain respectively. In this work, we consider the scenario in which an output feedback link is available from each receiver to its corresponding transmitter on top of the delayed knowledge of the channel state information as depicted in Figure 3(a).



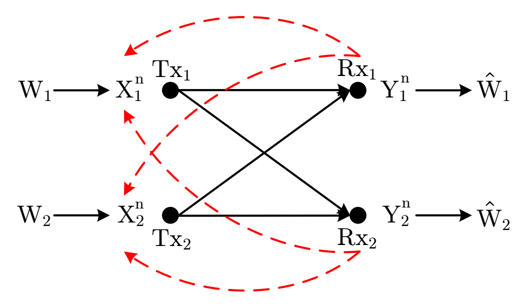


Fig. 3. *Two-user Binary Fading Interference Channel: (a) with output feedback links from each receiver to its corresponding transmitter. In this setting, the transmit signal of Tx_i at time instant t, would be a function of the message W_i, the available CSIT, and the output sequences Y_i^{t-1}, $i=1,2$; and (b) with output feedback links from each receiver to both transmitters. In this setting, the transmit signal of Tx_i at time instant t, would be a function of the message W_i, the available CSIT, and the output sequences Y_1^{t-1}, Y_2^{t-1}, $i=1,2$.*

In the presence of output feedback and Delayed-CSIT, we have the following result.

Theorem 3: **[Capacity Region with Delayed-CSIT and OFB]** For the two-user binary IC with Delayed-CSIT and OFB, the capacity region $\mathcal{C}^{\text{DCSIT,OFB}}$, is given by

$$\mathcal{C}^{\text{DCSIT,OFB}} = \left\{ R_1, R_2 \in \mathbb{R}^+ \ s.t. \ R_i + (1+q)R_{\bar{i}} \leq p(1+q)^2, \ i = 1,2 \right\}. \tag{12}$$

Remark 3: The outer-bound on the capacity region with only Delayed-CSIT (11) is in fact the intersection of the outer-bounds on the individual rates (*i.e.* $R_i \leq p$, $i = 1,2$) and the capacity region with Delayed-CSIT and OFB (12). Therefore, the impact of OFB is to remove the constraints on individual rates. This can be intuitively explained by noting that OFB creates new path to flow information from each transmitter to its corresponding receiver (*e.g.*, $\mathsf{Tx}_1 \to \mathsf{Rx}_2 \to \mathsf{Tx}_2 \to \mathsf{Rx}_1$). This opportunity results in elimination of the individual rate constraints in this case.

Remark 4: As we will see in Section VIII, same outer-bounds hold for the scenario in which output feedback links are available from each receiver to both transmitters, see Figure 3(b). Therefore, the capacity region of two-user binary IC with Delayed-CSIT and output feedback links are available from each receiver to both transmitters, is the same as the capacity region described in (12).

Finally, we present our result for the case of Instantaneous-CSIT and output feedback. Note that in this scenario, although transmitters have instantaneous knowledge of the channel state information, the output signals are available at the transmitters with unit delay. This scenario corresponds to a slow-fading channel where output feedback links are available from the receivers to the transmitters.

Theorem 4: **[Capacity Region with Instantaneous-CSIT and OFB]** For the two-user binary IC with Instantaneous-CSIT and OFB, the capacity region $\mathcal{C}^{\text{ICSIT,OFB}}$, is the set of all rate tuples (R_1, R_2) satisfying

$$\mathcal{C}^{\text{ICSIT,OFB}} = \begin{cases} 0 \leq R_i \leq 1 - q^2, & i = 1,2, \\ R_1 + R_2 \leq 1 - q^2 + pq. \end{cases} \tag{13}$$

Remark 5: Comparing the capacity region of the two-user BFIC with Instantaneous-CSIT, with OFB (13) or without OFB (10), we observe that the outer-bound on the sum-rate remains unchanged. However, the bounds on individual rates are further increased to $1 - q^2$. Similar to the previous remark, this is again due to the additional communication path provided by OFB from each transmitter to its intended receiver. However, since the outer-bound on sum-rate with Instantaneous-CSIT and OFB (13) is higher than that of Delayed-CSIT and OFB (12), the bounds on individual rates cannot be eliminated.

The proof of the results is organized as follows. The proof of Theorem 2 is presented in Sections V and VI. The proof of Theorem 3 is presented in Sections VII and VIII, and finally, the proof of Theorem 4 is presented in Sections IX and X. We end this section by illustrating our main results via an example in which $p = 0.5$.

C. Illustration of Main Results for $p = 0.5$

For this particular value of the channel parameter, the capacity region with Delayed-CSIT and Instantaneous-CSIT with or without output feedback is given in Table I, and Figure 4 illustrates the results presented in this table. We notice the following remarks.

TABLE I
ILLUSTRATION OF OUR MAIN RESULTS THROUGH AN EXAMPLE IN WHICH $p = 0.5$.

	Capacity Region with Delayed-CSIT	Capacity Region with Instantaneous-CSIT
No-OFB	$\begin{cases} R_i \leq \frac{1}{2} \\ R_i + \frac{3}{2}R_{\bar{i}} \leq \frac{9}{8} \end{cases}$	$\begin{cases} R_1 \leq \frac{1}{2} \\ R_2 \leq \frac{1}{2} \end{cases}$
OFB	$\begin{cases} R_1 + \frac{3}{2}R_2 \leq \frac{9}{8} \\ \frac{3}{2}R_1 + R_2 \leq \frac{9}{8} \end{cases}$	$\begin{cases} R_i \leq \frac{3}{4} \\ R_1 + R_2 \leq 1 \end{cases}$

Remark 6: Note that for $p = 0.5$, we have

$$\mathcal{C}^{\text{No}-\text{CSIT}} \subset \mathcal{C}^{\text{DCSIT}} \subset \mathcal{C}^{\text{ICSIT}}.$$

In other words, the capacity region with Instantaneous-CSIT is strictly larger than that of Delayed-CSIT, which is in turn strictly larger than the capacity region with No-CSIT. Moreover, we have

$$\mathcal{C}^{\text{DCSIT,OFB}} \subset \mathcal{C}^{\text{ICSIT,OFB}},$$

meaning that the instantaneous knowledge of the CSIT enlarges the capacity region of the two-user BFIC with OFB compared to the case of Delayed-CSIT.

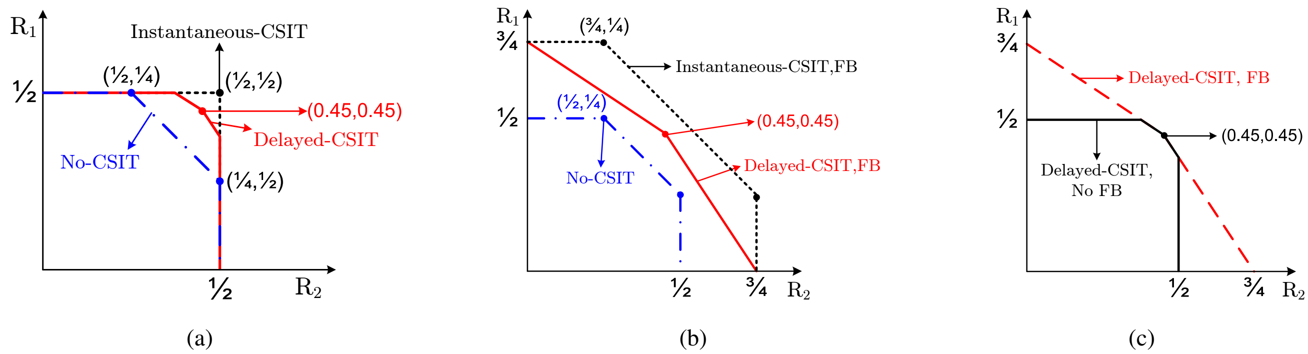


Fig. 4. *Two-user Binary Fading IC: (a) the capacity region with No-CSIT, Delayed-CSIT, and Instantaneous-CSIT, without OFB; (b) the capacity region with No-CSIT, Delayed-CSIT, and Instantaneous-CSIT, with OFB; and (c) the capacity region with Delayed-CSIT, with and without output feedback.*

Remark 7: In Figure 4(c), we have illustrated the capacity region with Delayed-CSIT, with and without output feedback. First, we observe that OFB enlarges the capacity region. Second, we observe that the optimal sum-rate point is the same for $p = 0.5$. However, this is not always the case. In fact, for some values of p, output feedback can even increase the optimal sum-rate. Using the results of Theorem 2 and Theorem 3, we have plotted the sum-rate capacity of the two-user Binary Fading IC with and without OFB for the Delayed-CSIT model in Figure 5. Note that for $0 < p < \left(3 - \sqrt{5}\right)/2$, the sum-rate capacity with OFB is strictly larger than the no OFB scenario.

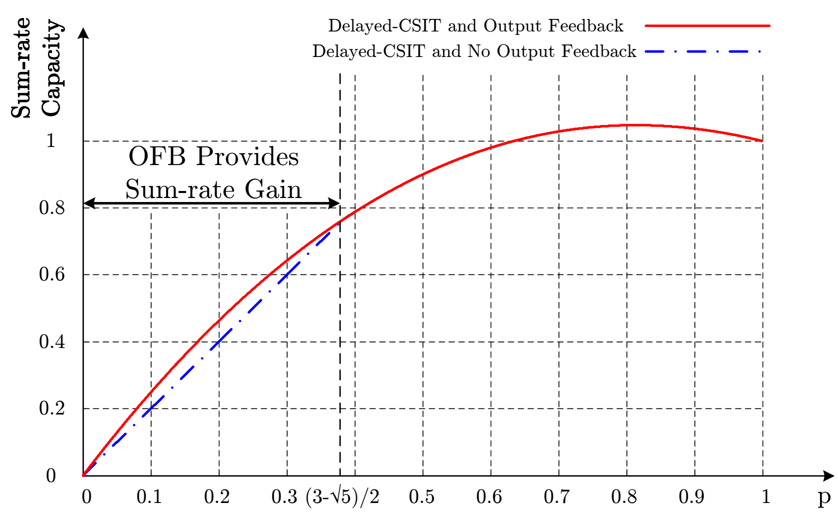


Fig. 5. The sum-rate capacity of the two-user BFIC with Delayed-CSIT, with and without output feedback. For $0 < p < \left(3 - \sqrt{5}\right)/2$, the sum-rate capacity with OFB is strictly larger than the scenario where no OFB is available.

Remark 8: Comparing the capacity region of the two-user BFIC with Instantaneous-CSIT, with OFB (13) or without OFB (10), we observe that OFB enlarges the capacity region. Moreover, similar to the Delayed-CSIT scenario, the optimal sum-rate point is the same for $p = 0.5$. Again, this is not always the case. In fact, for $0 < p < 0.5$, output feedback can even increase the optimal sum-rate. Using the results of Theorem 1 and Theorem 4, we have plotted the sum-rate capacity of the two-user Binary Fading IC with and without OFB in Figure 6.

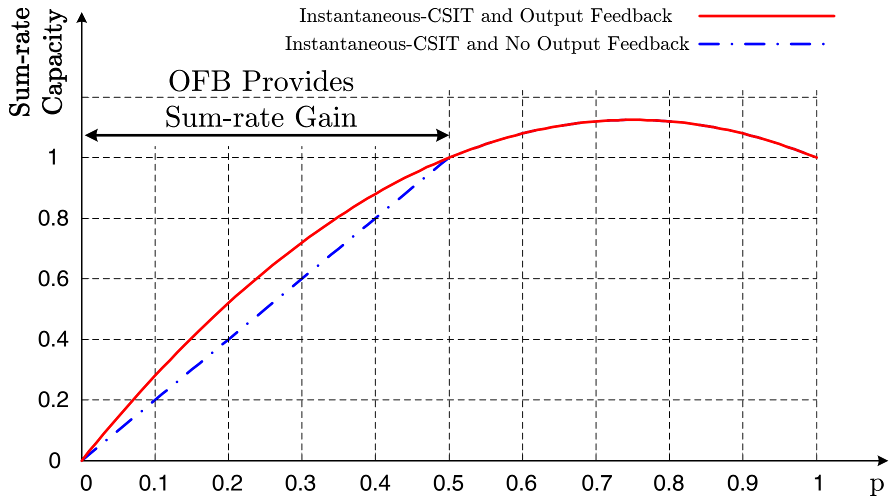


Fig. 6. The sum-rate capacity of the two-user BFIC with Instantaneous-CSIT, with and without output feedback. For $0 < p < 0.5$, the sum-rate capacity with OFB is strictly larger than the scenario where no OFB is available.

Remark 9: In Figures 5 and 6, we have identified the range of p for which output feedback provides sum-rate gain. Basically, when the sum-rate capacity without OFB is dominated by the capacity of the direct links (*i.e.* $2p$), and the additional communication paths created by the means of output feedback links help increase the optimal sum-rate.

In the following section, we present the main ideas that we incorporate in this paper.

IV. Overview of Key Ideas

Our goal in this section is to present the key techniques we use in this paper both for achievability and converse purposes. Although we will provide detailed explanation of the achievability strategy and converse proofs for all different scenarios, we found it instructive to elaborate the main ideas through several clarifying examples. In particular, for the sake of achievability, we show how Delayed-CSIT and later, Delayed-CSIT plus output feedback can help us improve the achievable rate region. Then, we present a key lemma that our converse proofs with Delayed-CSIT are based on.

A. Achievability Ideas with Delayed-CSIT

As we have described in Section II, the channel gains are independent from each other and over time. This way, transmitters cannot use the delayed knowledge of the channel state information to predict future. However, this information can still be very useful. In particular, Delayed-CSIT allows us to evaluate the contributions of the desired signal and the interference at each receiver in the past signaling stages and exploit it as available side information for future communication.

1) Interference-free Bits: Using Delayed-CSIT transmitters can identify previously transmitted bits such that if retransmitted, they do not create any further interference. The following examples clarify this idea.

Example 1 [Creating interference channels with side information]: Suppose at a time instant, each one of the transmitters simultaneously sends one data bit. The bits of Tx_1 and Tx_2 are denoted by a_1 and b_1 respectively. Later, using Delayed-CSIT, transmitters figure out that only the cross links were equal to 1 at this time instant as shown in Figure 7(a). This means that in future, transmission of these bits will no longer create interference at the unintended receivers. In other words, these bits can be transferred to a sub-problem, where in a two-user interference channel, Rx_i has apriori access to $\mathsf{W}_{\bar{i}}$, $i = 1, 2$. Since there will be no interference in this sub-problem, such bits can be communicated at higher rates.

Example 2 [Creating interference channels with swapped receivers and side information]: Assume that at a time instant, transmitters one and two simultaneously send data bits, say a_2 and b_2 respectively. Again through Delayed-CSIT, transmitters realize that all links except the link between Tx_1 and Rx_2 were equal to 1, see Figure 7(b).

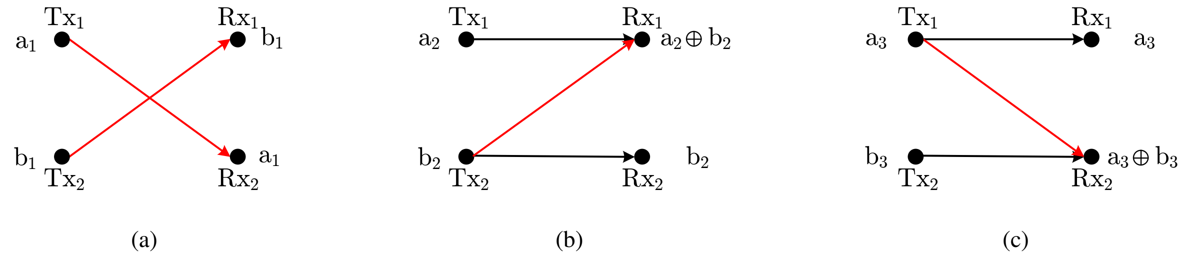


Fig. 7. *Achievability ideas with Delayed-CSIT:* (a) *via Delayed-CSIT transmitters can figure out which bits are already known at the unintended receivers, transmission of these bits will no longer create interference at the unintended receivers; and* (b) *to decode the bits, it is sufficient that* Tx_2 *provides* Rx_1 *with* b_2 *while this bit is available at* Rx_2*;* (c) *similar to* (b)*. Note that in* (b) *and* (c) *the intended receivers are swapped.*

In a similar case, assume that at another time instant, transmitters one and two send data bits a_3 and b_3 at the same time. Through Delayed-CSIT, transmitters realize that all links except the link between Tx_2 and Rx_1 were connected, see Figure 7(c). Then it is easy to see that to successfully finish delivering these bits, it is enough that Tx_1 sends a_3 to Rx_2, while this bit is already available at Rx_1; and Tx_2 sends b_2 to Rx_1, while it is already available at Rx_2. Therefore, again this problem can be evaluated as an interference channel with side information, where the intended receivers are *swapped*.

Remark 10: As described in Examples 1 and 2, an interference free bit can be retransmitted without worrying about creating interference at the unintended receiver. For instance consider Example 1, since the bits at Tx_i are already available at $\mathsf{Rx}_{\bar{i}}$, the transmitters can transfer such bits to an interference channel with side information as depicted in Figure 8. We note that for this new problem, the capacity region with no, delayed, or instantaneous CSIT is the same.

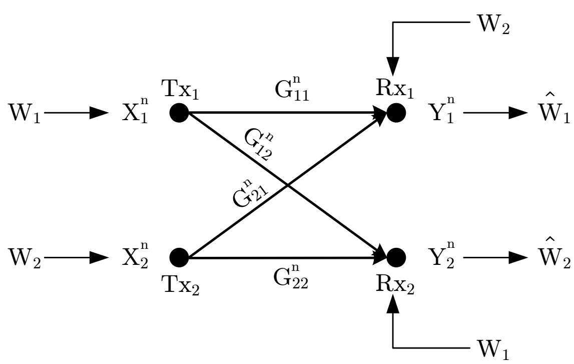


Fig. 8. Interference channel with side information: the capacity region with no, delayed, or instantaneous CSIT is the same.

2) Bits of Common Interest: Transmitters can use the delayed knowledge of the channel state information to identify bits that are of interest of both receivers. Below, we clarify this idea through several examples.

Example 3 [Opportunistic creation of bits of common interest]: Suppose at a time instant, each one of the transmitters sends one data bit, say a_4 and b_4 respectively. Later, using Delayed-CSIT, transmitters figure out that all links were equal to 1. In this case, both receivers have an equation of the transmitted bits, see Figure 9(a). Now, we notice that it is sufficient to provide either of the transmitted bits, a_4 or b_4, to both receivers rather than retransmitting both bits. We refer to such bits as bits of common interest. Since such bits are useful for both receivers, they can be transmitted more efficiently.

Remark 11 (Pairing bits of common interest to create a two-multicast problem): We note that in Example 3, one of the transmitters takes the responsibility of delivering one bit of common interest to both receivers. While this scheme improves the rate region, it is still not the most efficient scheme. The reason is that during the transmission, one of the transmitters stays silent and therefore the outgoing links of that transmitter are not utilized. To improve this, we can pair this problem with another similar problem as follows. Assume that in another time instant, each

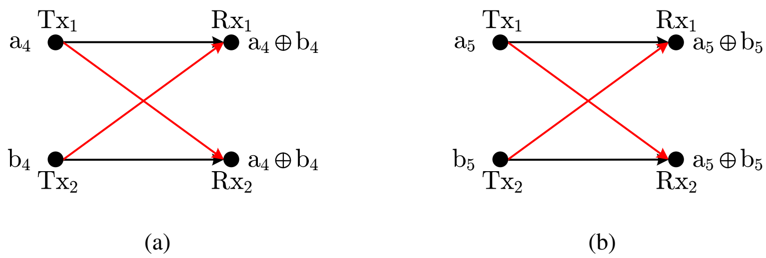


Fig. 9. *In each case, it is sufficient to provide only one of the transmitted bits to both receivers. We refer to such bits as bits of common interest.*

one of the transmitters sends one data bit, say a_5 and b_5 respectively. Later, transmitters figure out that all links were equal to 1, see Figure 9(b). In this case, similar to Example 3, one of the bits a_5 and b_5, say b_5, can be chosen as the bit of common interest. Now we can pair cases depicted in Figures 9(a) and 9(b). Then transmitters can simultaneously send bits a_4 and b_5, to both receivers. In this case, both bits a_4 and b_5 are needed by both receivers. In addition, we take advantage of all four links to transmit information, which improves the rate region.

Remark 12 (Pairing ICs with side information to create a two-multicast problem (pairing Type-I)): The advantage of interference channels with side information, explained in Examples 1 and 2, is that due to the side information, there is no interference involved in the problem. The downside is that half of the links in the channel become irrelevant and unexploited. More precisely, the cross links in Example 1 and the direct links in Example 2 are not utilized to increase the rate. Here we show that these two problems can be paired together to form an efficient two-multicast problem through creating bits of common interest. Referring to Figure 7, one can easily verify that it is enough to deliver $a_1 \oplus a_3$ and $b_1 \oplus b_2$ to both receivers. For instance, if $a_1 \oplus a_3$ and $b_1 \oplus b_2$ are available at Rx_1, it can remove b_1 from $b_1 \oplus b_2$ to decode b_2, then using b_2 and $a_2 \oplus b_2$ it can decode a_2; finally, using a_3 and $a_1 \oplus a_3$ it can decode a_1. Indeed, bit $a_1 \oplus a_3$ available at Tx_1, and bit $b_1 \oplus b_2$ available at Tx_2, are bits of common interest and can be transmitted to both receivers simultaneously in the efficient two-multicast problem as depicted in Figure 10. We note that for the two-multicast problem, the capacity region with no, delayed, or instantaneous CSIT is the same. We shall refer to this pairing as pairing **Type-I** throughout the paper.

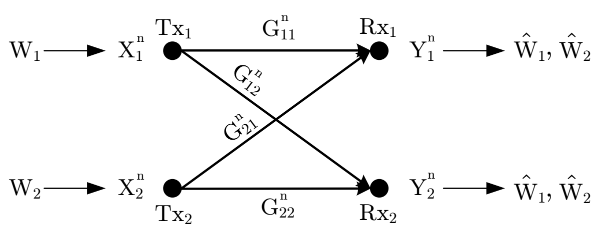


Fig. 10. Two-multicast network. Transmitter Tx_i wishes to reliably communicate message W_i to both receivers, $i = 1, 2$. The capacity region with no, delayed, or instantaneous CSIT is the same.

Example 4 [Pairing interference-free bits with bits of common interest to create a two-multicast problem (pairing Type-II)]: Suppose at a time instant, each one of the transmitters sends one data bit, say a_6 and b_6 respectively. Later, using Delayed-CSIT, transmitters figure out that all links were equal to 1, see Figure 11(a). In another time instant, each one of the transmitters sends one data bit, say a_7 and b_7 respectively. Later, transmitters figure out that only the cross links were equal to 1, see Figure 11(b). Now, we observe that providing $a_6 \oplus a_7$ and $b_6 \oplus b_7$ to both receivers is sufficient to decode the bits. For instance if Rx_1 is provided with $a_6 \oplus a_7$ and $b_6 \oplus b_7$, then it will use b_7 to decode b_6, from which it can obtain a_6, and finally using a_6 and $a_6 \oplus a_7$, it can decode a_7. Thus, bit $a_6 \oplus a_7$ available at Tx_1, and bit $b_6 \oplus b_7$ available at Tx_2, are bits of common interest and can be transmitted to both receivers simultaneously in the efficient two-multicast problem. We shall refer to this pairing as pairing **Type-II** throughout the paper.

Example 5 [Pairing bits of common interest with interference-free bits with swapped receivers to create a two-

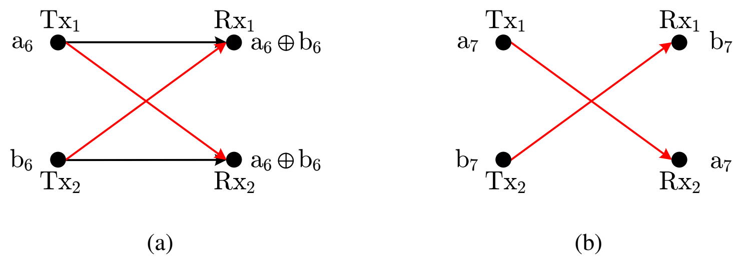


Fig. 11. *Pairing Type-II: providing $a_6 \oplus a_7$ and $b_6 \oplus b_7$ to both receivers is sufficient to decode the bits. In other words, bit $a_6 \oplus a_7$ available at Tx_1, and bit $b_6 \oplus b_7$ available at Tx_2, are bits of common interest and can be transmitted to both receivers simultaneously in the efficient two-multicast problem. Note that in (b), the cross links would have been irrelevant for future communications, however, using this pairing, we exploit all links.*

multicast problem (pairing Type-III)]: Suppose at a time instant, each one of the transmitters sends one data bit, say a_8 and b_8 respectively. Later, using Delayed-CSIT, transmitters figure out that all links were equal to 1 as in Figure 12(a). In another time instant, each one of transmitters sends one data bit, say a_9 and b_9 respectively. Later, transmitters figure out that all links were equal to 1 except the link from Tx_2 to Rx_1, see Figure 12(b). In a similar case, assume that at another time instant, transmitters one and two send data bits a_{10} and b_{10} at the same time. Through Delayed-CSIT, transmitters realize that all links except the link between Tx_1 and Rx_2 were connected, see Figure 12(c). We observe that providing $a_8 \oplus a_9$ and $b_8 \oplus b_{10}$ to both receivers is sufficient to decode the bits. For instance, if Rx_1 is provided with $a_8 \oplus a_9$ and $b_8 \oplus b_{10}$, then it will use a_9 to decode a_8, from which it can obtain b_8, then using b_8 and $b_8 \oplus b_{10}$, it gains access to b_{10}, finally using b_{10}, it can decode a_{10} from $a_{10} \oplus b_{10}$. Thus, bit $a_8 \oplus a_9$ available at Tx_1, and bit $b_8 \oplus b_{10}$ available at Tx_2, are bits of common interest and can be transmitted to both receivers simultaneously in the efficient two-multicast problem. We shall refer to this pairing as pairing **Type-III** throughout the paper.

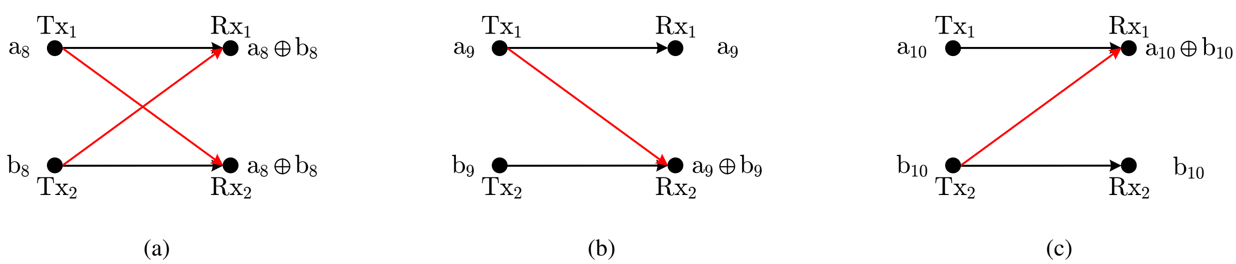


Fig. 12. *Pairing Type-III: providing $a_8 \oplus a_9$ and $b_8 \oplus b_{10}$ to both receivers is sufficient to decode the bits. In other words, bit $a_8 \oplus a_9$ available at Tx_1, and bit $b_8 \oplus b_{10}$ available at Tx_2, are bits of common interest and can be transmitted to both receivers simultaneously in the efficient two-multicast problem.*

As explained in the above examples, there are several ways to exploit the available side information at each transmitter. To achieve the capacity region, the first challenge is to evaluate various options and choose the most efficient one. The second challenge is that different opportunities may occur with different probabilities. This makes the process of matching, combining, and upgrading the status of the bits very difficult. Unfortunately, there is no simple guideline to decide when to search for the most efficient combination of the opportunities and when to hold on to other schemes. It is also important to note that most of the opportunities we observe here, do not appear in achieving the DoF of the Gaussian multi-antenna interference channels (see *e.g.*, [5], [6]).

B. Achievability Ideas with Output Feedback

In this subsection, we focus on the impact of the output feedback in the presence of Delayed-CSIT. The first observation is that through output feedback, each transmitter can evaluate the interference of the other transmitter, and therefore has access to the previously transmitted signal of the other user. Thus, output feedback can create

new path of communication between each transmitter and the corresponding receiver, *e.g.*,

$$\mathsf{Tx}_1 \rightarrow \mathsf{Rx}_2 \rightarrow \mathsf{Tx}_2 \rightarrow \mathsf{Rx}_1.$$

Although this additional path can improve the rate region, the advantage of output feedback is not limited to that. We explain the new opportunities through two examples

Example 6 [Creating two-multicast problem from ICs with side information]: In the previous subsection, we showed that interference-free transmissions can be upgraded to two-multicast problems through pairing. However, it is important to note that the different channel realizations used for pairing do not occur at the same probability. Therefore, it is not always possible to fully implement pairing in all cases. In particular, in some cases, some interference-free transmissions are left alone without possibility of pairing. In this example, we show that output feedback allows us to create bits of common interest out of these cases, which in turn allows us to create two-multicast problems. Referring to Figure 13, one can see that through the output feedback links, transmitters one and two can learn b_{11} and a_{11} respectively. Therefore, either of the transmitters is able to create $a_{11} \oplus b_{11}$. It is easy to see that $a_{11} \oplus b_{11}$ is of interest of both receivers. Indeed, feedback allows us to form a bit of common interest which can be delivered through the efficient two-multicast problem.

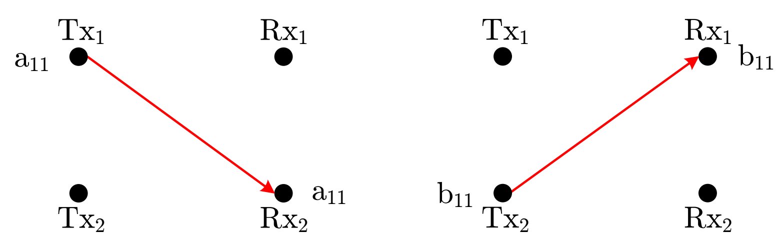


Fig. 13. *Opportunistic creation of bits of common interest using output feedback: bit b_{11} is available at Tx_1 via the feedback link from Rx_1; it is sufficient that Tx_1 provides $a_{11} \oplus b_{11}$ to both receivers.*

Example 7 [Creating two-multicast problem from ICs with swapped receivers and side information]: As another example, consider the two channel gain realizations depicted in Figure 14. In these cases, using output feedback Tx_1 can learn the transmitted bit of Tx_2 (*i.e.* b_{12}), and then form $a_{13} \oplus b_{12}$. It is easy to see that $a_{13} \oplus b_{12}$ is useful for both receivers and thus is a bit of common interest. Similar argument is valid for the second receiver. This means that output feedback allows us to upgrade interference-free transmissions with swapped receivers to bits of common interest that can be used to form efficient two-multicast problems.

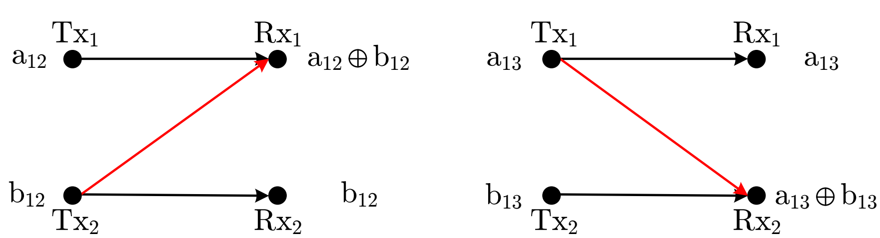


Fig. 14. *Opportunistic creation of bits of common interest using output feedback: using output feedback Tx_1 can learn the transmitted bit of Tx_2 (*i.e.* b_{12}); now, we observe that providing $a_{13} \oplus b_{12}$ to both receivers is sufficient to decode the intended bits.*

C. Key Idea for Converse Proofs with Delayed-CSIT

While we provide detailed proofs in Sections VI and VIII, we try to describe the main challenge in deriving the outer-bounds in this subsection. Consider the Delayed-CSIT scenario and suppose rate tuple (R_1, R_2) is achievable.

Then for $\beta > 0$, we have

$$
\begin{aligned}
n\left(R_1 + \beta R_2\right) &= H(W_1|W_2, G^n) + \beta H(W_2|G^n) \\
&\stackrel{\text{(Fano)}}{\leq} I(W_1; Y_1^n|W_2, G^n) + \beta I(W_2; Y_2^n|G^n) + n\epsilon_n \\
&= \beta H(Y_2^n|G^n) + \underbrace{H(G_{11}^n X_1^n|G^n) - \beta H(G_{12}^n X_1^n|G^n)} + n\epsilon_n.
\end{aligned} \tag{14}
$$

We refer the reader to Section VI for the detailed derivation of each step. Here, we would like to find a value of β such that

$$
H(G_{11}^n X_1^n|G^n) - \beta H(G_{12}^n X_1^n|G^n) \leq 0, \tag{15}
$$

for *any* input distribution. Note that since the terms involved are only a function of X_1^n and the channel gains, this term resembles a broadcast channel fromed by Tx_1 and the two receivers. Therefore, the main challenge boils down to understanding the ratio of the entropies of the received signals in a broadcast channel, and this would be the main focus of this subsection.

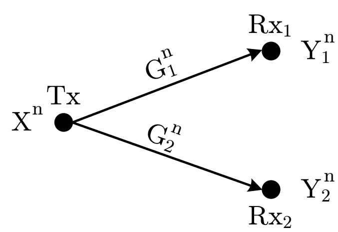


Fig. 15. A transmitter connected to two receivers through binary fading channels.

Consider a transmitter that is connected to two receivers through binary fading channels as depicted in Figure 15. We would like to understand how much this transmitter can privilege receiver one to receiver two, given outdated knowledge of the channel state information. Our metric would be the ratio of the entropies of the received signals[2]. In other words, we would like to understand what is the lower-bound on the ratio of the entropy of the received signal at Rx_2 to that of Rx_1. We first point out the result for the No-CSIT and Instantaneous-CSIT cases. With No-CSIT, from transmitter's point of view the two receivers are identical and it cannot favor one over the other and as a result, the two entropies would be equal. However with Instantaneous-CSIT, transmitter can choose to transmit at time t only if $G_1[t] = 1$ and $G_2[t] = 0$. Thus, with Instantaneous-CSIT the ratio of interest could be as low as 0. For the Delayed-CSIT case, we have the following lemma which we will formally prove in Section VI. Here, we try to provide some intuition about the problem by describing an input distribution that utilizes delayed knowledge of the channel state information in order to favor receiver one. It is important to keep in mind that this should not be considered as a proof but rather just a helpful intuition. Also, we point out that for the two-user BFIC with Delayed-CSIT and OFB, we will derive a variation of this lemma in Section VIII.

Lemma 1: **[Entropy Leakage]** For the channel described above with Delayed-CSIT, and for *any* input distribution, we have

$$
H\left(Y_2^n|G^n\right) \geq \frac{p}{1-q^2} H\left(Y_1^n|G^n\right). \tag{16}
$$

As mentioned before, we do not intend to prove this lemma here. We only provide an input distribution for which this lower-bound is tight. Consider m bits drawn from i.i.d. Bernoulli 0.5 random variables and assume these bits are in some initial queue. At any time instant t, the transmitter sends one of the bits in this initial queue (if the queue is empty, then the scheme is terminated). At time instant $t+1$, using Delayed-CSIT, the transmitter knows which one of the four possible channel realizations depicted in Figure 16 has occurred at time t. If either of the realizations (a) or (b) occurred at time t, then we remove the transmitted bit from the initial queue. However, if either of the realizations (c) or (d) occurred at time t, we leave this bit in the initial queue (*i.e.* among the bits that

[2]We point out that if $H(G_{11}^n X_1^n|G^n) = 0$, then ratio is not defined. But we keep in mind that what we really care about is (15).

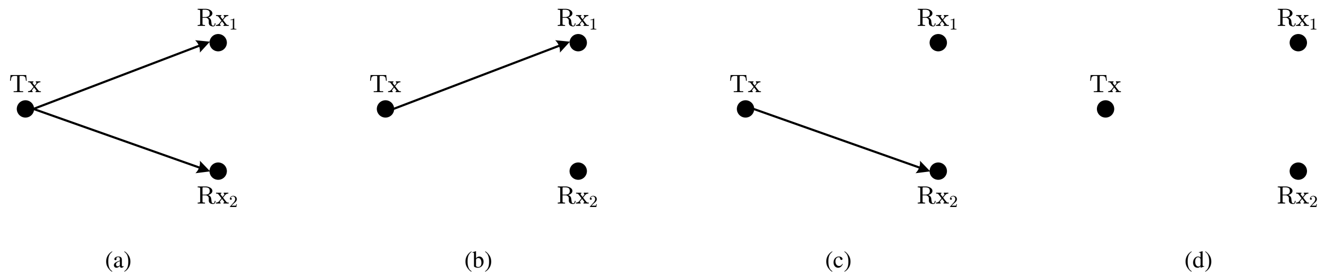


Fig. 16. *Four possible channel realizations for the network in Figure 15. The transmitter sends out a data bit at time instant t, and at the next time instant, using Delayed-CSIT, he knows which channel realization has occurred. If either of the realizations* (a) *or* (b) *occurred at time t, then we remove the transmitted bit from the initial queue. However, if either of the realizations* (c) *or* (d) *occurred at time t, we leave this bit in the initial queue. This way the transmitter favors receiver one over receiver two.*

can be transmitted at any future time instant). Note that this way, any bit that is available at Rx_2 would be available at Rx_1, however, there will be bits that are only available at Rx_1. Hence, the transmitter has favored receiver one over receiver two. If we analyze this scheme, we get

$$H\left(Y_2^n|G^n\right) = \frac{p}{1-q^2} H\left(Y_1^n|G^n\right), \tag{17}$$

meaning that the bound given in (16) is achievable and thus, it is tight.

Now that we have described the key ideas we incorporate in this paper, starting next section, we provide the proof of our main results.

V. Achievability Proof of Theorem 2 [Delayed-CSIT]

For $0 \leq p \leq \left(3-\sqrt{5}\right)/2$, the capacity of the two-user BFIC with Delayed-CSIT is depicted in Figure 17(a) and as a result, it is sufficient to describe the achievability for point $A = (p, p)$. However, for $\left(3-\sqrt{5}\right)/2 < p \leq 1$, all bounds are active and the region, as depicted in Figure 17(b), is the convex hull of points $A, B,$ and C. By symmetry, it is sufficient to describe the achievability for points A and C in this regime.

We first provide the achievability proof of point A for $0.5 \leq p \leq 1$ in this section. Then, we provide an overview of the achievability proof of corner point C and we postpone the detailed proof to Appendix D. Finally in Appendix C, we present the achievability proof of point A for $0 \leq p < 0.5$.

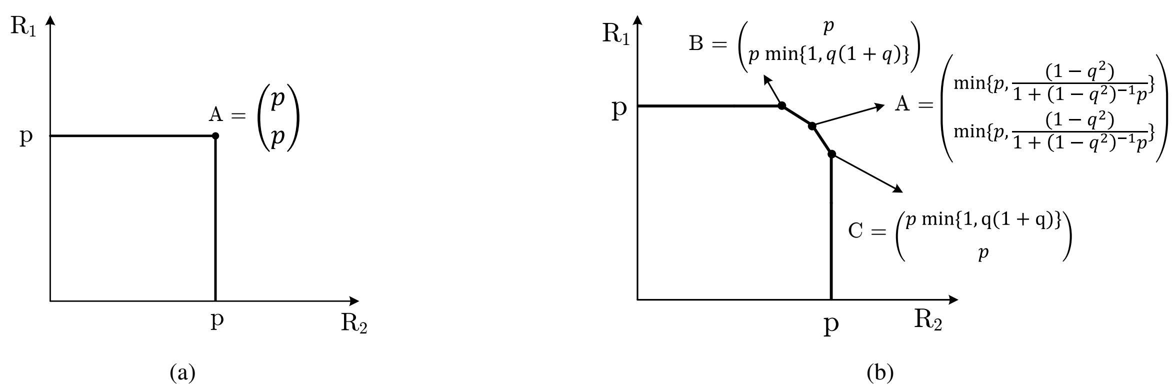
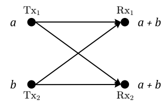
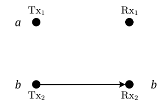
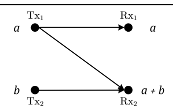
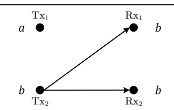
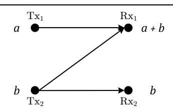
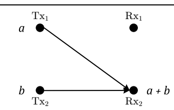
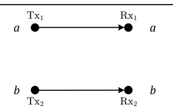
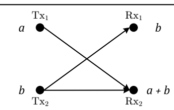
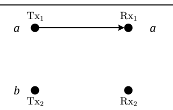
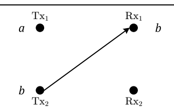
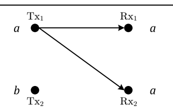
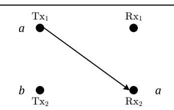
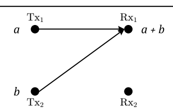
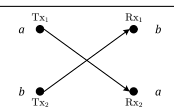
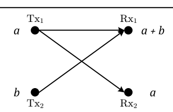
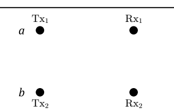


Fig. 17. *Capacity Region of the two-user Binary Fading IC with Delayed-CSIT for:* (a) $0 \leq p \leq \left(3-\sqrt{5}\right)/2$; *and* (b) $\left(3-\sqrt{5}\right)/2 < p \leq 1$.

TABLE II

ALL POSSIBLE CHANNEL REALIZATIONS AND TRANSITIONS FROM THE INITIAL QUEUE TO OTHER QUEUES; SOLID ARROW FROM TRANMSITTER Tx_i TO RECEIVER Rx_j INDICATES THAT $G_{ij}[t]=1$, $i,j\in\{1,2\}$, $t=1,2,\ldots,n$. BIT "a" REPRESENTS A BIT IN $Q_{1\to1}$ WHILE BIT "b" REPRESENTS A BIT IN $Q_{2\to2}$.

case ID	channel realization at time instant n	state transition	case ID	channel realization at time instant n	state transition
1		$\begin{cases} a\to Q_{1,\mathsf{C}_1} \\ b\to Q_{2,\mathsf{C}_1}\end{cases}$	9		$\begin{cases} a\to Q_{1\to1} \\ b\to Q_{2\to F}\end{cases}$
2		$\begin{cases} a\to Q_{1\to2\|1} \\ b\to Q_{2,F}\end{cases}$	10		$\begin{cases} a\to Q_{1\to1} \\ b\to Q_{2\to F}\end{cases}$
3		$\begin{cases} a\to Q_{1,F} \\ b\to Q_{2\to1\|2}\end{cases}$	11		$\begin{cases} a\to Q_{1\to\{1,2\}} \\ b\to Q_{2,F}\end{cases}$
4		$\begin{cases} a\to Q_{1\to F} \\ b\to Q_{2\to F}\end{cases}$	12		$\begin{cases} a\to Q_{1\to\{1,2\}} \\ b\to Q_{2,F}\end{cases}$
5		$\begin{cases} a\to Q_{1\to F} \\ b\to Q_{2\to2}\end{cases}$	13		$\begin{cases} a\to Q_{1\to1} \\ b\to Q_{2\to2\|1}\end{cases}$
6		$\begin{cases} a\to Q_{1\to F} \\ b\to Q_{2\to2}\end{cases}$	14		$\begin{cases} a\to Q_{1\to1\|2} \\ b\to Q_{2\to2}\end{cases}$
7		$\begin{cases} a\to Q_{1,F} \\ b\to Q_{2\to\{1,2\}}\end{cases}$	15		$\begin{cases} a\to Q_{1\to1\|2} \\ b\to Q_{2\to2\|1}\end{cases}$
8		$\begin{cases} a\to Q_{1,F} \\ b\to Q_{2\to\{1,2\}}\end{cases}$	16		$\begin{cases} a\to Q_{1\to1} \\ b\to Q_{2\to2}\end{cases}$

A. Achievability Strategy for Corner Point A

In this subsection, we describe a transmission strategy that achieves a rate tuple arbitrary close to corner point A for $0.5\leq p\leq 1$ as depicted in Figure 17(b), *i.e.*

$$R_1=R_2=\frac{(1-q^2)}{1+(1-q^2)^{-1}p}. \tag{18}$$

Let the messages of transmitters one and two be denoted by $\mathrm{W}_1=a_1,a_2,\ldots,a_m$, and $\mathrm{W}_2=b_1,b_2,\ldots,b_m$, respectively, where data bits a_i's and b_i's are picked uniformly and independently from $\{0,1\}$, $i=1,\ldots,m$. We show that it is possible to communicate these bits in

$$n=\left(1-q^2\right)^{-1}m+\left(1-q^2\right)^{-2}pm+O\left(m^{2/3}\right) \tag{19}$$

time instants[3] with vanishing error probability (as $m\to\infty$). Therefore achieving the rates given in (18) as $m\to\infty$. Our transmission strategy consists of two phases as described below.

Phase 1 [uncategorized transmission]: At the beginning of the communication block, we assume that the bits at Tx_i are in queue $Q_{i\to i}$ (the initial state of the bits), $i=1,2$. At each time instant t, Tx_i sends out a bit from $Q_{i\to i}$,

[3]Throughout the paper whenever we state the number of bits or time instants, say n, if the expression for a given value of p is not an integer, then we use the ceiling of that number $\lceil n\rceil$, where $\lceil.\rceil$ is the smallest integer greater than or equal to n. Note that since we will take the limit as $m\to\infty$, this does not change the end results.

and this bit will either stay in the initial queue or transition to one of the following possible queues will take place according to the description in Table II. If at time instant t, $Q_{i\to i}$ is empty, then Tx_i, $i = 1, 2$, remains silent until the end of Phase 1.

(A) Q_{i,C_1}: The bits that at the time of communication, all channel gains were equal to 1.
(B) $Q_{i\to\{1,2\}}$: The bits that are of common interest of both receivers and do not fall in category (A).
(C) $Q_{i\to i|\bar{i}}$: The bits that are required by Rx_i but are available at the unintended receiver $\mathsf{Rx}_{\bar{i}}$. A bit is in $Q_{i\to i|\bar{i}}$ if $\mathsf{Rx}_{\bar{i}}$ gets it without interference and Rx_i does not get it with or without interference.
(D) $Q_{i\to\bar{i}|i}$: The bits that are required by $\mathsf{Rx}_{\bar{i}}$ but are available at the intended receiver Rx_i. More precisely, a bit is in $Q_{i\to\bar{i}|i}$ if Rx_i gets the bit without interference and $\mathsf{Rx}_{\bar{i}}$ gets it with interference.
(E) $Q_{i\to F}$: The bits that we consider delivered and no retransmission is required.

More precisely, based on the channel realizations, a total of 16 possible configurations may occur at any time instant as summarized in Table II. The transition for each one of the channel realizations is as follows.

- Case 1 ( ): If at time instant t, Case 1 occurs, then each receiver gets a linear combination of the bits that were transmitted. Then as illustrated in Figure 18, if either of such bits is provided to both receivers then the receivers can decode both bits. The transmitted bit of Tx_i leaves $Q_{i\to i}$ and joins Q_{i,C_1}[4], $i = 1, 2$. Although we can consider such bits as bits of common interest, we keep them in an intermediate queue for now and as we describe later, we combine them with other bits to create bits of common interest.

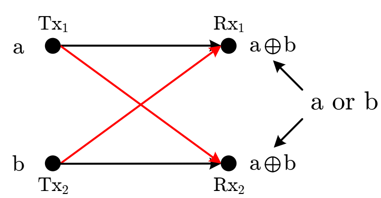


Fig. 18. Suppose transmitters one and two send out data bits a and b respectively and Case 1 occurs. Now, if either of the transmitted bits is provided to both receivers, then each receiver can decode its corresponding bit.

- Case 2 ( ): In this case, Rx_1 has already received its corresponding bit while Rx_2 has a linear combination of the transmitted bits, see Table II. As a result, if the transmitted bit of Tx_1 is provided to Rx_2, it will be able to decode both bits. In other words, the transmitted bit from Tx_1 is available at Rx_1 and is required by Rx_2. Therefore, transmitted bit of Tx_1 leaves $Q_{1\to 1}$ and joins $Q_{1\to 2|1}$. Note that the bit of Tx_2 will not be retransmitted since upon delivery of the bit of Tx_1, Rx_2 can decode its corresponding bit. Since no retransmission is required, the bit of Tx_2 leaves $Q_{2\to 2}$ and joins $Q_{2,F}$ (the final state of the bits).
- Case 3 ( ): This is similar to Case 2 with swapping user IDs.
- Case 4 ( ): In this case, each receiver gets its corresponding bit without any interference. We consider such bits to be delivered and no retransmission is required. Therefore, the transmitted bit of Tx_i leaves $Q_{i\to i}$ and joins $Q_{i,F}$, $i = 1, 2$.
- Case 5 ( ) and Case 6 ( ): In these cases, Rx_1 gets its corresponding bit interference free. We consider this bit to be delivered and no retransmission is required. Therefore, the transmitted bit of Tx_1 leaves $Q_{1\to 1}$ and joins $Q_{1,F}$, while the transmitted bit of Tx_2 remains in $Q_{2\to 2}$.
- Case 7 ( ): In this case, Rx_1 has a linear combination of the transmitted bits, while Rx_2 has not received anything, see Table II. It is sufficient to provide the transmitted bit of Tx_2 to both receivers. Therefore, the transmitted bit of Tx_2 leaves $Q_{2\to 2}$ and joins $Q_{2\to\{1,2\}}$. Note that the bit of Tx_1 will not be retransmitted since upon delivery of the bit of Tx_2, Rx_1 can decode its corresponding bit. This bit leaves $Q_{1\to 1}$ and joins $Q_{1,F}$. Similar argument holds for Case 8 ( ).

[4]In this paper, we assume that the queues are ordered. Meaning that the first bit that joins the queue is placed at the head of the queue and any new bit occupies the next empty position. For instance, suppose there are ℓ bits in Q_{1,C_1} and ℓ bits in Q_{2,C_1}, then the next time Case 1 occurs, the transmitted bit of Tx_i is placed at position $\ell + 1$ in Q_{i,C_1}, $i = 1, 2$.

- Cases 9,10,11, and 12: Similar to Cases 5,6,7, and 8 with swapping user IDs respectively.
- Case 13 ($\nearrow$): In this case, Rx_1 has received the transmitted bit of Tx_2 while Rx_2 has not received anything, see Table II. Therefore, the transmitted bit of Tx_1 remains in $Q_{1\rightarrow 1}$, while the transmitted bit of Tx_2 is required by Rx_2 and it is available at Rx_1. Hence, the transmitted bit of Tx_2 leaves $Q_{2\rightarrow 2}$ and joins $Q_{2\rightarrow 2|1}$. Queue $Q_{2\rightarrow 2|1}$ represents the bits at Tx_2 that are available at Rx_1, but Rx_2 needs them.
- Case 14 ($\searrow$): This is similar to Case 13 with swapping user IDs.
- Case 15 ($\nearrow\!\!\!\!\searrow$): In this case, Rx_1 has received the transmitted bit of Tx_2 while Rx_2 has received the transmitted bit of Tx_1, see Table II. In other words, the transmitted bit of Tx_2 is available at Rx_1 and is required by Rx_2; while the transmitted bit of Tx_1 is available at Rx_2 and is required by Rx_1. Therefore, we have transition from $Q_{i\rightarrow i}$ to $Q_{i\rightarrow i|\bar{i}}$, $i=1,2$.
- Case 16: The transmitted bit of Tx_i remains in $Q_{i\rightarrow i}$, $i=1,2$.

Phase 1 goes on for

$$\left(1-q^2\right)^{-1} m + m^{\frac{2}{3}} \tag{20}$$

time instants, and if at the end of this phase, either of the queues $Q_{i\rightarrow i}$ is not empty, we declare error type-I and halt the transmission (we assume m is chosen such that $m^{\frac{2}{3}} \in \mathbb{Z}$).

Assuming that the transmission is not halted, let N_{i,C_1}, $N_{i\rightarrow j|\bar{j}}$, and $N_{i\rightarrow \{1,2\}}$ denote the number of bits in queues Q_{i,C_1}, $Q_{i\rightarrow j|\bar{j}}$, and $Q_{i\rightarrow \{1,2\}}$ respectively at the end of the transitions, $i=1,2$, and $j=i,\bar{i}$. The transmission strategy will be halted and an error type-II will occur, if any of the following events happens.

$$\begin{aligned}
&N_{i,\mathsf{C}_1} > \mathbb{E}[N_{i,\mathsf{C}_1}] + m^{\frac{2}{3}} \overset{\triangle}{=} n_{i,\mathsf{C}_1}, \quad i=1,2;\\
&N_{i\rightarrow j|\bar{j}} > \mathbb{E}[N_{i\rightarrow j|\bar{j}}] + m^{\frac{2}{3}} \overset{\triangle}{=} n_{i\rightarrow j|\bar{j}}, \quad i=1,2, \text{ and } j=i,\bar{i};\\
&N_{i\rightarrow \{1,2\}} > \mathbb{E}[N_{i\rightarrow \{1,2\}}] + m^{\frac{2}{3}} \overset{\triangle}{=} n_{i\rightarrow \{1,2\}}, \quad i=1,2.
\end{aligned} \tag{21}$$

From basic probability, we know that

$$\begin{aligned}
\mathbb{E}[N_{i,\mathsf{C}_1}] &= \frac{\Pr\left(\text{Case 1}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m = (1-q^2)^{-1}p^4 m,\\
\mathbb{E}[N_{i\rightarrow i|\bar{i}}] &= \frac{\sum_{j=14,15}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m = (1-q^2)^{-1}pq^2 m,\\
\mathbb{E}[N_{i\rightarrow \bar{i}|i}] &= \frac{\Pr\left(\text{Case 2}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m = (1-q^2)^{-1}p^3 q m,\\
\mathbb{E}[N_{i\rightarrow \{1,2\}}] &= \frac{\sum_{j=11,12}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m = (1-q^2)^{-1}p^2 q m.
\end{aligned} \tag{22}$$

Furthermore, we can show that the probability of errors of types I and II decreases exponentially with m. More precisely, we use Chernoff-Hoeffding bound[5], to bound the error probabilities of types I and II. For instance, to bound the probability of error type-I, we have

$$\begin{aligned}
\Pr\left[\text{error type}-\text{I}\right] &\leq \sum_{i=1}^{2} \Pr\left[Q_{i\rightarrow i} \text{ is not empty}\right]\\
&\leq 4\exp\left(\frac{-m^{4/3}}{4n(1-q^2)q^2}\right)\\
&= 4\exp\left(\frac{-m^{4/3}}{4(1-q^2)q^2\left[\left(1-q^2\right)^{-1}m + m^{\frac{2}{3}}\right]}\right),
\end{aligned} \tag{23}$$

[5]We consider a specific form of the Chernoff-Hoeffding bound [21] described in [22], which is simpler to use and is as follows. If $X_1,\ldots,X_r$ are r independent random variables, and $M=\sum_{i=1}^{r} X_i$, then $Pr\left[|M-\mathbb{E}\left[M\right]| > \alpha\right] \leq 2\exp\left(\frac{-\alpha^2}{4\sum_{i=1}^{r}\text{Var}(X_i)}\right)$.

which decreases exponentially to zero as $m \to \infty$.

At the end of Phase 1, we add 0's (if necessary) in order to make queues Q_{i,C_1}, $Q_{i \to j|\bar{j}}$, and $Q_{i \to \{1,2\}}$ of size equal to n_{i,C_1}, $n_{i \to j|\bar{j}}$, and $n_{i \to \{1,2\}}$ respectively as defined in (21), $i = 1, 2$, and $j = i, \bar{i}$. For the rest of this subsection, we assume that Phase 1 is completed and no error has occurred.

We now use the ideas described in Section IV-A, to further create bits of common interest. Depending on the value of p, we use different ideas. We break the rest of this subsection into two parts: (1) $0.5 \leq p \leq \left(\sqrt{5}-1\right)/2$; and (2) $\left(\sqrt{5}-1\right)/2 < p \leq 1$. In what follows, we first describe the rest of the achievability strategy for $0.5 \leq p \leq \left(\sqrt{5}-1\right)/2$. In particular, we demonstrate how to incorporate the ideas of Section IV-A to create bits of common interest in an optimal way.

- **Type I** Combining bits in $Q_{i \to \bar{i}|i}$ and $Q_{i \to i|\bar{i}}$: Consider the bits that were transmitted in Cases 2 and 14, see Figure 19. Observe that if we provide $a_1 \oplus a_2$ to *both* receivers then Rx_1 can decode bits a_1 and a_2, whereas Rx_2 can decode bit b_1. Therefore, $a_1 \oplus a_2$ is a bit of common interest and can join $Q_{1 \to \{1,2\}}$. Hence, as illustrated in Figure 20, we can remove two bits in $Q_{1 \to 2|1}$ and $Q_{1 \to 1|2}$, by inserting their XOR in $Q_{1 \to \{1,2\}}$, and we deliver this bit of common interest to both receivers during the second phase. Note that due to the symmetry of the channel, similar argument holds for $Q_{2 \to 1|2}$ and $Q_{2 \to 2|1}$.

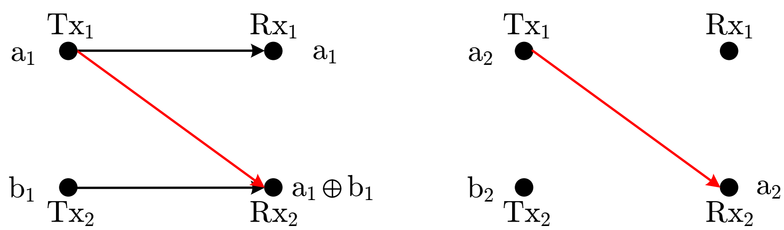


Fig. 19. Suppose at a time instant, transmitters one and two send out data bits a_1 and b_1 respectively, and later using Delayed-CSIT, transmitters figure out Case 2 occurred at that time. At another time instant, suppose transmitters one and two send out data bits a_2 and b_2 respectively, and later using Delayed-CSIT, transmitters figure out Case 14 occurred at that time. Now, bit $a_1 \oplus a_2$ available at Tx_1 is useful for both receivers and it is a bit of common interest. Hence, $a_1 \oplus a_2$ can join $Q_{1 \to \{1,2\}}$.

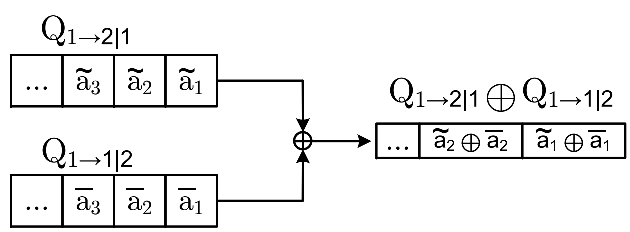


Fig. 20. Creating XOR of the bits in two different queues. We pick one bit from each queue and create the XOR of the two bits.

For $0.5 \leq p \leq \left(\sqrt{5}-1\right)/2$, we have $\mathbb{E}[N_{i \to \bar{i}|i}] \leq \mathbb{E}[N_{i \to i|\bar{i}}]$. Therefore, after this combination, queue $Q_{i \to \bar{i}|i}$ becomes empty and we have

$$\mathbb{E}[N_{i \to i|\bar{i}}] - \mathbb{E}[N_{i \to \bar{i}|i}] = pq\left(q - p^2\right) \tag{24}$$

bits left in $Q_{i \to i|\bar{i}}$.

- **Type II** Combining the bits in Q_{i,C_1} and $Q_{i \to i|\bar{i}}$: Consider the bits that were transmitted in Cases 1 and 15, see Figure 21. It is easy to see that providing $a_1 \oplus a_2$ and $b_1 \oplus b_2$ to both receivers is sufficient to decode their corresponding bits. For instance, Rx_1 removes b_2 from $b_1 \oplus b_2$ to decode b_1, then uses b_1 to decode a_1 from $a_1 \oplus b_1$. Therefore, $a_1 \oplus a_2$ and $b_1 \oplus b_2$ are bits of common interest and can join $Q_{1 \to \{1,2\}}$ and $Q_{2 \to \{1,2\}}$ respectively. Hence, we can remove two bits in Q_{i,C_1} and $Q_{i \to i|\bar{i}}$, by inserting their XOR in $Q_{i \to \{1,2\}}$, $i = 1, 2$, and then deliver this bit of common interest to both receivers during the second phase.

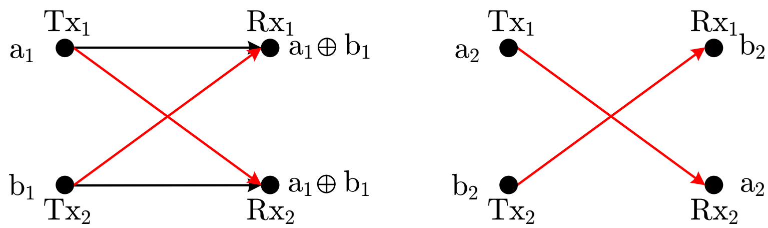


Fig. 21. Suppose at a time instant, transmitters one and two send out data bits a_1 and b_1 respectively, and later using Delayed-CSIT, transmitters figure out Case 1 occurred at that time. At another time instant, suppose transmitters one and two send out data bits a_2 and b_2 respectively, and later using Delayed-CSIT, transmitters figure out Case 15 occurred at that time. Now, bit $a_1 \oplus a_2$ available at Tx_1 and bit $b_1 \oplus b_2$ available at Tx_2 are useful for both receivers and they are bits of common interest. Therefore, bits $a_1 \oplus a_2$ and $b_1 \oplus b_2$ can join $Q_{1\rightarrow\{1,2\}}$ and $Q_{2\rightarrow\{1,2\}}$ respectively.

For $0.5 \leq p \leq \left(\sqrt{5}-1\right)/2$, we have $pq\left(q-p^2\right) \leq \mathbb{E}[N_{i,\mathsf{C}_1}]$. Therefore after combining the bits, queue $Q_{i\rightarrow i|\bar{i}}$ becomes empty and we have

$$\mathbb{E}[N_{i,\mathsf{C}_1}] + m^{\frac{2}{3}} - \left(\mathbb{E}[N_{i\rightarrow i|\bar{i}}] - \mathbb{E}[N_{i\rightarrow \bar{i}|i}]\right) = p\left(p-q\right) + m^{\frac{2}{3}} \tag{25}$$

bits left in Q_{i,C_1}, $i=1,2$.

Finally, we need to describe what happens to the remaining $p\left(p-q\right)+m^{\frac{2}{3}}$ bits in Q_{i,C_1}. As mentioned before, a bit in Q_{i,C_1} can be viewed as a bit of common interest by itself. For the remaining bits in Q_{1,C_1}, we put the first half in $Q_{1\rightarrow\{1,2\}}$ (suppose m is picked such that the remaining number of bits is even). Note that if these bits are delivered to Rx_2, then Rx_2 can decode the first half of the remaining bits in Q_{2,C_1} as well. Therefore, the first half of the bits in Q_{2,C_1} can join $Q_{2,F}$.

Then, we put the second half of the remaining bits in Q_{2,C_1} in $Q_{2\rightarrow\{1,2\}}$. Similar to the argument presented above, the second half of the bits in Q_{1,C_1} join $Q_{1,F}$.

Hence at the end of Phase 1, if the transmission is not halted, we have a total of

$$\begin{aligned}(1-q^2)^{-1}&\left[\underbrace{p^2q+m^{2/3}}_{\text{Cases 11 and 12}} + \underbrace{pq^2+m^{2/3}}_{\text{XOR opportunities}} + 0.5\underbrace{\left(p^4-pq^2+p^3q+m^{2/3}\right)}_{\text{remaining Case 1}}\right]m \\ &= \left(1-q^2\right)^{-1}0.5pm + 2.5m^{2/3}\end{aligned} \tag{26}$$

number of bits in $Q_{1\rightarrow\{1,2\}}$. Same result holds for $Q_{2\rightarrow\{1,2\}}$.

This completes the description of Phase 1 for $0.5 \leq p \leq \left(\sqrt{5}-1\right)/2$. For $\left(\sqrt{5}-1\right)/2 < p \leq 1$, we combine the bits as follows. Fot this range of p, after Phase 1, the number of bits in each queue is such that the mergings described above are not optimal and we have to rearrange them as decsribed below.

- **Type I** Combining $Q_{i\rightarrow \bar{i}|i}$ and $Q_{i\rightarrow i|\bar{i}}$: We have already described this opportunity for $0.5 \leq p \leq \left(\sqrt{5}-1\right)/2$. We create the XOR of the bits in $Q_{1\rightarrow 2|1}$ and $Q_{1\rightarrow 1|2}$ and put the XOR of them in $Q_{1\rightarrow\{1,2\}}$. Note that due to the symmetry of the channel, similar argument holds for $Q_{2\rightarrow 1|2}$ and $Q_{2\rightarrow 2|1}$.
  For $\left(\sqrt{5}-1\right)/2 < p \leq 1$, $\mathbb{E}[N_{i\rightarrow i|\bar{i}}] \leq \mathbb{E}[N_{i\rightarrow \bar{i}|i}]$, $i=1,2$. Therefore after combining the bits, queue $Q_{i\rightarrow i|\bar{i}}$ becomes empty, and we have
  $$\mathbb{E}[N_{i\rightarrow \bar{i}|i}] - \mathbb{E}[N_{i\rightarrow i|\bar{i}}] = pq\left(p^2-q\right) \tag{27}$$
  bits left in $Q_{i\rightarrow \bar{i}|i}$, $i=1,2$.
- **Type III** Combining the bits in Q_{i,C_1} and $Q_{i\rightarrow \bar{i}|i}$: Consider the bits that were transmitted in Cases $1,2$, and 3, see Figure 22. Now, we observe that providing $a\oplus c$ and $b \oplus f$ to both receivers is sufficient to decode their corresponding bits. For instance, Rx_1 will have $a\oplus b$, c, $e\oplus f$, $a\oplus c$, and $b\oplus f$, from which it can recover a, c, and e. Similar argument holds for Rx_2. Therefore, $a\oplus c$ and $b\oplus f$ are bits of common interest and can join $Q_{1\rightarrow\{1,2\}}$ and $Q_{2\rightarrow\{1,2\}}$ respectively. Hence, we can remove two bits in Q_{i,C_1} and $Q_{i\rightarrow \bar{i}|i}$, by inserting their XORs in $Q_{i\rightarrow\{1,2\}}$, $i=1,2$, and then deliver this bit of common interest to both receivers during the second phase.

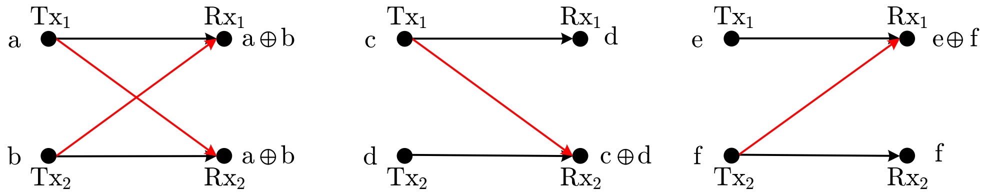


Fig. 22. Consider the bits transmitted in Cases 1,2, and 3. Now, bit $a \oplus c$ available at Tx_1 and bit $b \oplus f$ available at Tx_2 are useful for both receivers and they are bits of common interest. Therefore, bits $a \oplus c$ and $b \oplus f$ can join $Q_{1\rightarrow\{1,2\}}$ and $Q_{2\rightarrow\{1,2\}}$ respectively.

For $\left(\sqrt{5}-1\right)/2 < p \leq 1$, we have $pq\left(p^2-q\right) \leq \mathbb{E}[N_{i,\mathsf{C}_1}]$. Therefore after combining the bits, queue $Q_{i\rightarrow\bar{i}|i}$ becomes empty and we have

$$\mathbb{E}[N_{i,\mathsf{C}_1}] + m^{\frac{2}{3}} - \left(\mathbb{E}[N_{i\rightarrow\bar{i}|i}] - \mathbb{E}[N_{i\rightarrow i|\bar{i}}]\right) = p^4 - p^3q + pq^2 + m^{\frac{2}{3}} \tag{28}$$

bits left in Q_{i,C_1}.

We treat the remaining bits in Q_{i,C_1} as described before. Hence at the end of Phase 1, if the transmission is not halted, we have a total of

$$\begin{aligned}&(1-q^2)^{-1}\left[\underbrace{p^2q+m^{2/3}}_{\text{Cases 11 and 12}} + \underbrace{p^3q+m^{2/3}}_{\text{XOR opportunities}} + 0.5\underbrace{\left(p^4-p^3q+pq^2+m^{2/3}\right)}_{\text{remaining Case 1}}\right]m \\ &= \left(1-q^2\right)^{-1}0.5pm + 2.5m^{2/3}\end{aligned} \tag{29}$$

number of bits in $Q_{1\rightarrow\{1,2\}}$. Same result holds for $Q_{2\rightarrow\{1,2\}}$.

To summarize, at the end of Phase 1 assuming that the transmission is not halted, by using coding opportunities of types I, II, and III, we are only left with $\left(1-q^2\right)^{-1}0.5pm + 2.5m^{2/3}$ bits in queue $Q_{i\rightarrow\{1,2\}}$, $i = 1, 2$.

We now describe how to deliver the bits of common interest in Phase 2 of the transmission strategy. The problem resembles a network with two transmitters and two receivers where each transmitter Tx_i wishes to communicate an independent message W_i to *both* receivers, $i = 1, 2$. The channel gain model is the same as described in Section II. We refer to this network as the two-multicast network as depicted in Figure 23. We have the following result for this network.

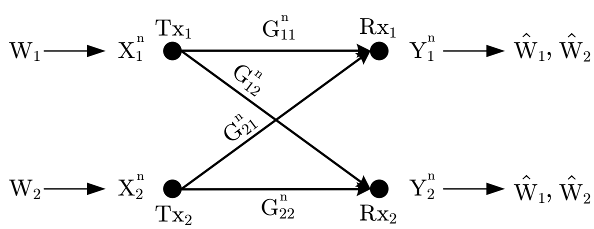


Fig. 23. Two-multicast network. Transmitter Tx_i wishes to reliably communicate message W_i to both receivers, $i = 1, 2$. The capacity region with no, delayed, or instantaneous CSIT is the same.

Lemma 2: For the two-multicast network as described above, we have

$$\mathcal{C}^{\text{No-CSIT}}_{\text{multicast}} = \mathcal{C}^{\text{DCSIT}}_{\text{multicast}} = \mathcal{C}^{\text{ICSIT}}_{\text{multicast}}, \tag{30}$$

and, we have

$$\mathcal{C}^{\text{ICSIT}}_{\text{multicast}} = \begin{cases} R_i \leq p, & i = 1, 2, \\ R_1 + R_2 \leq 1 - q^2. & \end{cases} \tag{31}$$

This result basically shows that the capacity region of the two-multicast network described above is equal to the capacity region of the multiple-access channel formed at either of the receivers. The proof of Lemma 2 is presented in Appendix E.

Phase 2 [transmitting bits of common interest]: In this phase, we deliver the bits in $Q_{1\to\{1,2\}}$ and $Q_{2\to\{1,2\}}$ using the transmission strategy for the two-multicast problem. More precisely, the bits in $Q_{i\to\{1,2\}}$ will be considered as the message of Tx_i and they will be encoded as in the achievability scheme of Lemma 2, $i = 1, 2$. Fix $\epsilon, \delta > 0$, from Lemma 2, we know that rate tuple

$$(R_1, R_2) = \frac{1}{2}\left((1-q^2) - \delta, (1-q^2) - \delta\right)$$

is achievable with decoding error probability less than or equal to ϵ. Therefore, transmission of the bits in $Q_{1\to\{1,2\}}$ and $Q_{2\to\{1,2\}}$, will take

$$t_{\text{total}} = \frac{\left(1-q^2\right)^{-1} pm + 5m^{2/3}}{(1-q^2) - \delta}. \tag{32}$$

Therefore, the total transmssion time of our two-phase achievability strategy is equal to

$$(1-q^2)^{-1}m + m^{\frac{2}{3}} + \frac{\left(1-q^2\right)^{-1} pm + 5m^{2/3}}{(1-q^2) - \delta}, \tag{33}$$

hence, if we let $\epsilon, \delta \to 0$ and $m \to \infty$, the decoding error probability of delivering bits of common interest goes to zero, and we achieve a symmetric sum-rate of

$$R_1 = R_2 = \lim_{\substack{\epsilon,\delta\to 0\\ m\to\infty}} \frac{m}{t_{\text{total}}} = \frac{(1-q^2)}{1+(1-q^2)^{-1}p}. \tag{34}$$

This completes the achievability proof of point A for $0.5 \leq p \leq 1$.

B. Overview of the Achievability Strategy for Corner Point C

We now provide an overview of the achievability strategy for corner point C depicted in Figure 17 for $\left(3-\sqrt{5}\right)/2 < p \leq 1$, *i.e.*

$$(R_1, R_2) = \left(pq(1+q), p\right), \tag{35}$$

and we postpone the detailed proof to Appendix D.

Compared to the achievability strategy of the sum-rate point, the challenges in achieving the other corner points arise from the asymmetricity of the rates. At this corner point, while Tx_2 (the primary user) communicates at full rate of p, Tx_1 (the secondary user) communicates at a lower rate and tries to coexist with the primary user. The achievability strategy is based on the following two principles.

1) If the secondary user creates interference at the primary receiver, it is the secondary user's responsibility to resolve this interference;
2) For the achievability of the optimal sum-rate point A (see Figure 17(b)), the bits of common interest were transmitted such that both receivers could decode them. However, for corner point C, when the primary receiver obtains a bit of common interest, we revise the coding scheme in a way that favors the primary receiver.

Our transmission strategy consists of five phases as summarized below.

- Phase 1 [uncategorized transmission]: This phase is similar to Phase 1 of the achievability of the optimal sum-rate point A. The main difference is due to the fact that the transmitters have unequal number of bits at the beginning. In Phase 1, Tx_1 (the secondary user) transmits all its initial bits while Tx_2 (the primary user) only transmits part of its initial bits. Transmitter two postpones the transmission of its remaining bits to Phase 3.
- Phase 2 [updating status of the bits transmitted when either of the Cases 7 or 8 (11 or 12) occurred]: For the achievability of optimal sum-rate point A, we transferred the transmitted bits of Tx_2 (Tx_1) to the two-multicast sub-problem by viewing them as bits of common interest. However, this scheme turns out to be suboptimal

for corner point C. In this case, we retransmit these bits during Phase 2 and update their status based on the channel realization at the time of transmission. Phase 2 provides coding opportunities that we exploit in Phases 4 and 5.

- Phase 3 [uncategorized transmission vs interference management]: In this phase, the primary user transmits the remaining initial bits while the secondary user tries to resolve as much interference as it can at the primary receiver. To do so, the secondary user sends the bits that caused interference at the primary receiver during Phase 1, at a rate low enough such that both receivers can decode and remove them regardless of what the primary transmitter does. Note that pq of the time, each receiver gets interference-free signal from the secondary transmitter, hence, the secondary transmitter can take advantage of these time instants to deliver its bits during Phase 3.
- Phases 4 and 5 [delivering interference-free bits and interference management]: In the final phases, each transmitter has two main objectives: (1) communicating the bits required by its own receiver but available at the unintended receiver; and (2) mitigating interference at the unintended receiver. This task can be accomplished by creating the XOR of the bits similar to coding type-I described in Section IV with a modification: we first encode these bits and then create the XOR of the encoded bits. Moreover, the balance of the two objectives is different between the primary user and the secondary user.

As mentioned before, the detailed proof of the achievability for corner point C is provided in Appendix D. In the following section, we describe the converse proof for the two-user BFIC with Delayed-CSIT.

VI. Converse Proof of Theorem 2 [Delayed-CSIT]

In this section, we provide the converse proof for Theorem 2. We first present a lemma that plays a key role in deriving the converse.

Consider the scenario where a transmitter is connected to two receivers through binary fading channels as in Figure 24. Suppose $G_1[t]$ and $G_2[t]$ are distributed as i.i.d. Bernoulli RVs (*i.e.* $G_i[t] \overset{d}{\sim} \mathcal{B}(p)$), $i = 1, 2$. In this channel the received signals are given as

$$Y_i[t] = G_i[t]X[t], \qquad i = 1, 2, \tag{36}$$

where $X[t]$ is the transmit signal at time t. We have the following lemma.

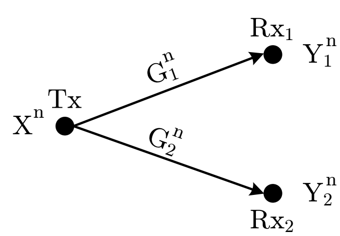


Fig. 24. A transmitter connected to two receivers through binary fading channels. Using Delayed-CSIT, the transmitter can privilege reveiver one to receiver two. Lemma 3 formalizes this privilege.

Lemma 3: **[Entropy Leakage]** For the channel described above with Delayed-CSIT and for *any* input distribution, we have

$$H\left(Y_2^n|G^n\right) \geq \frac{1}{2-p} H\left(Y_1^n|G^n\right). \tag{37}$$

Remark 13: Note that with No-CSIT, from the transmitter's point of view, the two receivers are identical and it cannot favor one over the other and as a result, we have $H\left(Y_2^n|G^n\right) = H\left(Y_1^n|G^n\right)$. With Instantaneous-CSIT this ratio can become zero[6]. Therefore, this lemma captures the impact of Delayed-CSIT on the entropy of the received signals at the two receivers.

[6]This can be done by simply remaining silent whenever $G_2[t] = 1$.

Proof: For time instant t where $1 \leq t \leq n$, we have

$$\begin{aligned}
&H\left(Y_2[t]|Y_2^{t-1},G^t\right)\\
&\quad= pH\left(X[t]|Y_2^{t-1},G_2[t]=1,G^{t-1}\right)\\
&\quad\overset{(a)}{=} pH\left(X[t]|Y_2^{t-1},G^t\right)\\
&\quad\overset{(b)}{\geq} pH\left(X[t]|Y_1^{t-1},Y_2^{t-1},G^t\right)\\
&\quad\overset{(c)}{=} \frac{p}{1-q^2}H\left(Y_1[t],Y_2[t]|Y_1^{t-1},Y_2^{t-1},G^t\right),
\end{aligned} \tag{38}$$

where (a) holds since $X[t]$ is independent of the channel realization at time instant t; (b) follows from the fact that conditioning reduces entropy; and (c) follows from the fact that $\Pr\left[G_1[t]=G_2[t]=0\right]=q^2$. Therefore, we have

$$\sum_{t=1}^{n} H\left(Y_2[t]|Y_2^{t-1},G^t\right) \geq \frac{1}{2-p}\sum_{t=1}^{n} H\left(Y_1[t],Y_2[t]|Y_1^{t-1},Y_2^{t-1},G^t\right), \tag{39}$$

and since the transmit signals at time instant t are independent from the channel realizations in future time instants, we have

$$\sum_{t=1}^{n} H\left(Y_2[t]|Y_2^{t-1},G^n\right) \geq \frac{1}{2-p}\sum_{t=1}^{n} H\left(Y_1[t],Y_2[t]|Y_1^{t-1},Y_2^{t-1},G^n\right), \tag{40}$$

hence, we get

$$H\left(Y_2^n|G^n\right) \geq \frac{1}{2-p}H\left(Y_1^n,Y_2^n|G^n\right) \geq \frac{1}{2-p}H\left(Y_1^n|G^n\right). \tag{41}$$

This completes the proof of the lemma. ∎

We now derive the converse for Theorem 2. The outer-bound on R_i is the same under no, delayed, and instantaneous CSIT, and we present it in Appendix B. In this section, we provide the proof of

$$R_i + (1+q)R_{\bar{i}} \leq p(1+q)^2, \qquad i=1,2. \tag{42}$$

By symmetry, it is sufficient to prove it for $i=1$. Let $\beta=(1+q)$, and suppose rate tuple (R_1,R_2) is achievable. Then we have

$$\begin{aligned}
&n\left(R_1+\beta R_2\right) = H(W_1)+\beta H(W_2)\\
&\quad\overset{(a)}{=} H(W_1|W_2,G^n)+\beta H(W_2|G^n)\\
&\quad\overset{\text{(Fano)}}{\leq} I(W_1;Y_1^n|W_2,G^n)+\beta I(W_2;Y_2^n|G^n)+n\epsilon_n\\
&\quad= H(Y_1^n|W_2,G^n) - \underbrace{H(Y_1^n|W_1,W_2,G^n)}_{=\,0} + \beta H(Y_2^n|G^n) - \beta H(Y_2^n|W_2,G^n)+n\epsilon_n\\
&\quad\overset{(b)}{=} \beta H(Y_2^n|G^n) + H(Y_1^n|W_2,X_2^n,G^n) - \beta H(Y_2^n|W_2,X_2^n,G^n)+n\epsilon_n\\
&\quad= \beta H(Y_2^n|G^n) + H(G_{11}^nX_1^n|W_2,X_2^n,G^n) - \beta H(G_{12}^nX_1^n|W_2,X_2^n,G^n)+n\epsilon_n\\
&\quad\overset{(c)}{=} \beta H(Y_2^n|G^n) + H(G_{11}^nX_1^n|W_2,G^n) - \beta H(G_{12}^nX_1^n|W_2,G^n)+n\epsilon_n\\
&\quad\overset{(d)}{=} \beta H(Y_2^n|G^n) + H(G_{11}^nX_1^n|G^n) - \beta H(G_{12}^nX_1^n|G^n)+n\epsilon_n\\
&\quad\overset{\text{Lemma 3}}{\leq} \beta H(Y_2^n|G^n)+n\epsilon_n\\
&\quad= \beta\sum_{t=1}^{n} H(Y_2[t]|Y_2^{t-1},G^n)+n\epsilon_n\\
&\quad\overset{(e)}{\leq} \beta\sum_{t=1}^{n} H(Y_2[t]|G^n)+n\epsilon_n\\
&\quad\overset{(f)}{\leq} \beta(1-q^2)n+\epsilon_n = p(1+q)^2n+\epsilon_n.
\end{aligned} \tag{43}$$

where (a) holds since W_1, W_2 and G^n are mutually independent; (b) and (c) hold since X_2^n is a deterministic function of W_2 and G^n; (d) follows from

$$\begin{aligned} 0 &\leq H(G_{11}^n X_1^n | G^n) - H(G_{11}^n X_1^n | W_2, G^n) \\ &= I\left(G_{11}^n X_1^n ; W_2 | G^n\right) \leq I\left(W_1, G_{11}^n X_1^n ; W_2 | G^n\right) \\ &= \underbrace{I\left(W_1 ; W_2 | G^n\right)}_{=\,0 \text{ since } \mathrm{W}_1 \perp \mathrm{W}_2 \perp \mathrm{G}^n} + \underbrace{I\left(G_{11}^n X_1^n ; W_2 | W_1, G^n\right)}_{=\,0 \text{ since } \mathrm{X}_1 = \mathrm{f}_1(\mathrm{W}_1,\, \mathrm{G}^n)} = 0, \end{aligned} \tag{44}$$

which implies $H(G_{11}^n X_1^n | G^n) = H(G_{11}^n X_1^n | W_2, G^n)$, and similarly $H(G_{12}^n X_1^n | G^n) = H(G_{12}^n X_1^n | W_2, G^n)$; (e) is true since conditioning reduces entropy; and (f) holds since the probability that at least one of the links connected to Rx_2 is equal to 1 at each time instant is $(1-q^2)$. Dividing both sides by n and let $n \to \infty$, we get

$$R_1 + (1+q) R_2 \leq p(1+q)^2. \tag{45}$$

This completes the converse proof for Theorem 2.

VII. Achievability Proof of Theorem 3 [Delayed-CSIT and OFB]

We now focus on the impact of the output feedback in the presence of Delayed-CSIT. In particular, we demonstrate how output feedback can be utilized to further improve the achievable rates. The capacity region of the two-user BFIC with Delayed-CSIT and OFB is given by

$$\mathcal{C}^{\mathrm{DCSIT,OFB}} = \left\{ R_1, R_2 \in \mathbb{R}^+ \ s.t. \ R_i + (1+q) R_{\bar{i}} \leq p(1+q)^2, \ i = 1, 2 \right\}, \tag{46}$$

and is depicted in Figure 25.

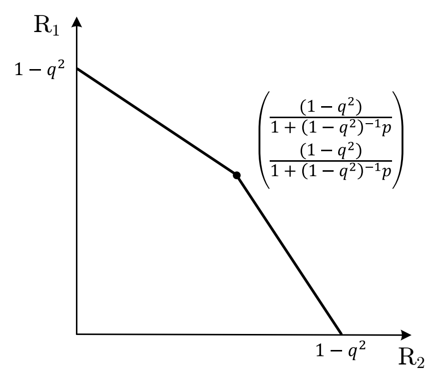


Fig. 25. Capacity region of the two-user BFIC with Delayed-CSIT and output feedback.

The achievability strategy of the corner points $\left(1-q^2, 0\right)$ and $\left(0, 1-q^2\right)$, is based on utilizing the additional communication paths created by the means of the output feedback links, *e.g.*,

$$\mathsf{Tx}_1 \to \mathsf{Rx}_2 \to \mathsf{Tx}_2 \to \mathsf{Rx}_1,$$

and is presented in Appendix F. Here, we only describe the transmission strategy for the sum-rate point, *i.e.*

$$R_1 = R_2 = \frac{(1-q^2)}{1 + (1-q^2)^{-1} p}. \tag{47}$$

Let the messages of transmitters one and two be denoted by $\mathrm{W}_1 = a_1, a_2, \ldots, a_m$, and $\mathrm{W}_2 = b_1, b_2, \ldots, b_m$, respectively, where data bits a_i's and b_i's are picked uniformly and independently from $\{0,1\}$, $i = 1, \ldots, m$. We show that it is possible to communicate these bits in

$$n = \left(1-q^2\right)^{-1} m + \left(1-q^2\right)^{-2} pm + O\left(m^{2/3}\right) \tag{48}$$

time instants with vanishing error probability (as $m \to \infty$). Therefore achieving the rates given in (47) as $m \to \infty$. Our transmission strategy consists of two phases as described below.

Phase 1 [uncategorized transmission]: This phase is identical to Phase 1 of Section V. At the beginning of the communication block, we assume that the bits at Tx_i are in queue $Q_{i\rightarrow i}$, $i = 1, 2$. At each time instant t, Tx_i sends out a bit from $Q_{i\rightarrow i}$, and this bit will either stay in the initial queue or transition to a new queue will take place. The transitions are identical to what we have already described in Table II, therefore, we are not going to repeat them here. Phase 1 goes on for

$$\left(1-q^2\right)^{-1} m + m^{\frac{2}{3}} \tag{49}$$

time instants and if at the end of this phase, either of the queues $Q_{1\rightarrow 1}$ or $Q_{2\rightarrow 2}$ is not empty, we declare error type-I and halt the transmission.

The transmission strategy will be halted and an error type-II will occur, if any of the following events happens.

$$\begin{aligned}
&N_{i,\mathsf{C}_1} > \mathbb{E}[N_{i,\mathsf{C}_1}] + m^{\frac{2}{3}} \overset{\triangle}{=} n_{i,\mathsf{C}_1}, \quad i = 1,2;\\
&N_{i\rightarrow j|\bar{j}} > \mathbb{E}[N_{i\rightarrow j|\bar{j}}] + m^{\frac{2}{3}} \overset{\triangle}{=} n_{i\rightarrow j|\bar{j}}, \quad i = 1,2, \text{ and } j = i, \bar{i};\\
&N_{i\rightarrow \{1,2\}} > \mathbb{E}[N_{i\rightarrow \{1,2\}}] + m^{\frac{2}{3}} \overset{\triangle}{=} n_{i\rightarrow \{1,2\}}, \quad i = 1,2.
\end{aligned} \tag{50}$$

From basic probability, we know that

$$\begin{aligned}
\mathbb{E}[N_{i,\mathsf{C}_1}] &= \frac{\Pr\left(\text{Case 1}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m = (1-q^2)^{-1}p^4 m,\\
\mathbb{E}[N_{i\rightarrow i|\bar{i}}] &= \frac{\sum_{j=14,15}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m = (1-q^2)^{-1}pq^2 m,\\
\mathbb{E}[N_{i\rightarrow \bar{i}|i}] &= \frac{\Pr\left(\text{Case 2}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m = (1-q^2)^{-1}p^3 q m,\\
\mathbb{E}[N_{i\rightarrow \{1,2\}}] &= \frac{\sum_{j=11,12}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m = (1-q^2)^{-1}p^2 q m.
\end{aligned} \tag{51}$$

Using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II decreases exponentially with m.

At the end of Phase 1, we add 0's (if necessary) in order to make queues Q_{i,C_1}, $Q_{i\rightarrow j|\bar{j}}$, and $Q_{i\rightarrow \{1,2\}}$ of size equal to n_{i,C_1}, $n_{i\rightarrow j|\bar{j}}$, and $n_{i\rightarrow \{1,2\}}$ respectively as defined in (50), $i = 1, 2$, and $j = i, \bar{i}$. For the rest of this subsection, we assume that Phase 1 is completed and no error has occurred. We now use the ideas described in Section IV for output feedback, to further create bits of common interest.

• Updating the status of bits in Q_{i,C_1} to bits of common interest: As described in Example 3 of Section IV-B, a bit in Q_{i,C_1} can be considered as a bit of common interest. Also note that it is sufficient to deliver only one of the two bits transmitted simultaneously during Case 1. Therefore, Tx_1 updates the status of the first half of the bits in Q_{1,C_1} to $Q_{1\rightarrow \{1,2\}}$, whereas Tx_2 updates the status of the second half of the bits in Q_{2,C_1} to $Q_{2\rightarrow \{1,2\}}$. Hence, after updating the status of bits in Q_{i,C_1}, we have

$$(1-q^2)^{-1}\left[p^2 q + \frac{1}{2}p^4\right] m + \frac{3}{2} m^{\frac{2}{3}} \tag{52}$$

bits in $Q_{i\rightarrow \{1,2\}}$, $i = 1, 2$.

• Upgrading ICs with side information to a two-multicast problem using OFB: Note that through the output feedback links, each transmitter has access to the transmitted bits of the other user during Phase 1. As described above, there are $\mathbb{E}[N_{i\rightarrow i|\bar{i}}] + m^{\frac{2}{3}}$ bits in $Q_{i\rightarrow i|\bar{i}}$ at the end of Phase 1. Now, Tx_1 creates the XOR of the first half of the bits in $Q_{1\rightarrow 1|2}$ and $Q_{2\rightarrow 2|1}$ and updates the status of the resulting bits to $Q_{1\rightarrow \{1,2\}}$. Note that as described in Example 6 of Section IV, the XOR of these bits is a bit of common interest. On the other hand, Tx_2 creates the XOR of the second half of the bits in $Q_{1\rightarrow 1|2}$ and $Q_{2\rightarrow 2|1}$ and updates the status of the resulting bits to $Q_{2\rightarrow \{1,2\}}$. Thus, we have

$$(1-q^2)^{-1}\left[p^2 q + \frac{1}{2}p^4 + \frac{1}{2}pq^2\right] m + 2m^{\frac{2}{3}} \tag{53}$$

bits in $Q_{i\to\{1,2\}}$, $i=1,2$.

• Upgrading ICs with side information and swapped receivers to a two-multicast problem using OFB: As described above, there are $\mathbb{E}[N_{i\to\bar{i}|i}]+m^{\frac{2}{3}}$ bits in $Q_{i\to\bar{i}|i}$, Tx_1 creates the XOR of the first half of the bits in $Q_{1\to 2|1}$ and $Q_{2\to 1|2}$ and updates the status of the resulting bits to $Q_{1\to\{1,2\}}$. Note that as described in Example 7 of Section IV, the XOR of these bits is a bit of common interest. On the other hand, Tx_2 creates the XOR of the second half of the bits in $Q_{1\to 2|1}$ and $Q_{2\to 1|2}$ and updates the status of the resulting bits to $Q_{2\to\{1,2\}}$. Hence, we have

$$(1-q^2)^{-1}\left[p^2q+\frac{1}{2}p^4+\frac{1}{2}pq^2+\frac{1}{2}p^3q\right]m+\frac{5}{2}m^{\frac{2}{3}}=(1-q^2)^{-1}\frac{p}{2}m+\frac{5}{2}m^{\frac{2}{3}} \tag{54}$$

bits in $Q_{i\to\{1,2\}}$, $i=1,2$. This completes the description of Phase 1.

Phase 2 [transmitting bits of common interest]: In this phase, we deliver the bits in $Q_{1\to\{1,2\}}$ and $Q_{2\to\{1,2\}}$ using the transmission strategy for the two-multicast problem. More precisely, the bits in $Q_{i\to\{1,2\}}$ will be considered as the message of Tx_i and they will be encoded as in the achievability scheme of Lemma 2, $i=1,2$. Fix $\epsilon,\delta>0$, from Lemma 2 we know that the rate tuple

$$(R_1,R_2)=\frac{1}{2}\Big((1-q^2)-\delta/2,(1-q^2)-\delta/2\Big)$$

is achievable with decoding error probability less than or equal to ϵ. Therefore, transmission of the bits in $Q_{1\to\{1,2\}}$ and $Q_{2\to\{1,2\}}$ requires

$$t_{\text{total}}=\frac{\left(1-q^2\right)^{-1}pm+5m^{2/3}}{(1-q^2)-\delta} \tag{55}$$

time instants. Therefore, the total transmission time of our two-phase achievability strategy is equal to

$$(1-q^2)^{-1}m+m^{\frac{2}{3}}+\frac{\left(1-q^2\right)^{-1}pm+5m^{2/3}}{(1-q^2)-\delta}. \tag{56}$$

The probability that the transmission strategy halts at any point can be bounded by the summation of error probabilities of types I and II, and the probability that an error occurs in decoding the encoded bits. This probability approaches zero for $\epsilon,\delta\to 0$ and $m\to\infty$.

Hence, if we let $\epsilon,\delta\to 0$ and $m\to\infty$, the decoding error probability goes to zero, and we achieve a symmetric sum-rate of

$$R_1=R_2=\lim_{\substack{\epsilon,\delta\to 0\\ m\to\infty}}\frac{m}{t_{\text{total}}}=\frac{(1-q^2)}{1+(1-q^2)^{-1}p}. \tag{57}$$

VIII. Converse Proof of Theorem 3 [Delayed-CSIT and OFB]

In this section, we prove the converse for Theorem 3. Suppose rate tuple (R_1, R_2) is achievable, then by letting $\beta = 1+q$, we have

$$\begin{aligned}
n\left(R_1+\beta R_2\right) &= H(W_1)+\beta H(W_2)\\
&\overset{(a)}{=} H(W_1|W_2,G^n)+\beta H(W_2|G^n)\\
&\overset{\text{Fano}}{\leq} I(W_1;Y_1^n|W_2,G^n)+\beta I(W_2;Y_2^n|G^n)+n\epsilon_n\\
&\leq I(W_1;Y_1^n,Y_2^n|W_2,G^n)+\beta I(W_2;Y_2^n|G^n)+n\epsilon_n\\
&= H(Y_1^n,Y_2^n|W_2,G^n)-\underbrace{H(Y_1^n,Y_2^n|W_1,W_2,G^n)}_{=\,0}+\beta H(Y_2^n|G^n)-\beta H(Y_2^n|W_2,G^n)+n\epsilon_n\\
&= \beta H(Y_2^n|G^n)+H(Y_1^n,Y_2^n|W_2,G^n)-\beta H(Y_2^n|W_2,G^n)+n\epsilon_n\\
&= \beta H(Y_2^n|G^n)+\sum_{t=1}^{n}H(Y_1[t],Y_2[t]|W_2,Y_1^{t-1},Y_2^{t-1},G^n)-\beta\sum_{t=1}^{n}H(Y_2[t]|W_2,Y_2^{t-1},G^n)+n\epsilon_n\\
&\overset{(b)}{\leq} \beta H(Y_2^n|G^n)+\sum_{t=1}^{n}H(Y_1[t],Y_2[t]|W_2,Y_1^{t-1},Y_2^{t-1},X_2^t,G^n)-\beta\sum_{t=1}^{n}H(Y_2[t]|W_2,Y_2^{t-1},X_2^t,G^n)+n\epsilon_n\\
&= \beta H(Y_2^n|G^n)+\sum_{t=1}^{n}H(G_{11}[t]X_1[t],G_{12}[t]X_1[t]|W_2,G_{11}^{t-1}X_1^{t-1},G_{12}^{t-1}X_1^{t-1},X_2^t,G^n)\\
&\quad-\beta\sum_{t=1}^{n}H(G_{12}[t]X_1[t]|W_2,G_{12}^{t-1}X_1^{t-1},X_2^t,G^n)+n\epsilon_n\\
&\overset{(c)}{=} \beta H(Y_2^n|G^n)+\sum_{t=1}^{n}H(G_{11}[t]X_1[t],G_{12}[t]X_1[t]|W_2,G_{11}^{t-1}X_1^{t-1},G_{12}^{t-1}X_1^{t-1},X_2^t,G^t)\\
&\quad-\beta\sum_{t=1}^{n}H(G_{12}[t]X_1[t]|W_2,G_{12}^{t-1}X_1^{t-1},X_2^t,G^t)+n\epsilon_n\\
&\overset{(d)}{\leq} \beta H(Y_2^n|G^n)+n\epsilon_n\\
&\leq p(1+q)^2 n+n\epsilon_n,
\end{aligned}\tag{58}$$

where (a) holds since the channel gains and the messages are mutually independent; (b) follows from the fact that X_2^t is a deterministic function of $\left(W_2, Y_2^{t-1}\right)$[7] and the fact that conditioning reduces entropy; (c) follows from the fact that condition on W_2, X_1^{t-1}, X_2^t, $X_1[t]$ is independent of the channel realization at future time instants, hence, we can replace G^n by G^t; and (d) follows from Lemma 4 below. Dividing both sides by n and let $n\to\infty$, we get

$$R_1+(1+q)R_2 \leq p(1+q)^2. \tag{59}$$

Similarly, we can get $(1+q)R_1+R_2 \leq p(1+q)^2$.

Lemma 4:

$$\begin{aligned}
&\sum_{t=1}^{n}H(G_{12}[t]X_1[t]|W_2,G_{12}^{t-1}X_1^{t-1},X_2^t,G^t)\\
&\qquad\geq \frac{1}{2-p}\sum_{t=1}^{n}H(G_{11}[t]X_1[t],G_{12}[t]X_1[t]|W_2,G_{11}^{t-1}X_1^{t-1},G_{12}^{t-1}X_1^{t-1},X_2^t,G^t).
\end{aligned}\tag{60}$$

[7] We have also added Y_1^{t-1} in the condition for the scenario in which output feedback links are available from each receiver to both transmitters.

Proof: We have

$$\begin{aligned}
&H(G_{12}[t]X_1[t]|W_2, G_{12}^{t-1}X_1^{t-1}, X_2^t, G^t)\\
&= pH(X_1[t]|G_{12}[t]=1, W_2, G_{12}^{t-1}X_1^{t-1}, X_2^t, G^{t-1})\\
&\stackrel{(a)}{=} pH(X_1[t]|W_2, G_{12}^{t-1}X_1^{t-1}, X_2^t, G^{t-1})\\
&\stackrel{(b)}{=} pH(X_1[t]|W_2, G_{12}^{t-1}X_1^{t-1}, X_2^t, G^t)\\
&= \frac{p}{1-q^2}H(G_{11}[t]X_1[t], G_{12}[t]X_1[t]|W_2, G_{12}^{t-1}X_1^{t-1}, X_2^t, G^t)\\
&\stackrel{(c)}{\geq} \frac{1}{2-p}H(G_{11}[t]X_1[t], G_{12}[t]X_1[t]|W_2, G_{11}^{t-1}X_1^{t-1}, G_{12}^{t-1}X_1^{t-1}, X_2^t, G^t),
\end{aligned} \tag{61}$$

where (a) and (b) follow from the fact that condition on W_2, $G_{12}^{t-1}X_1^{t-1}$, X_2^t and G^n, $X_1[t]$ is independent of the channel realization at time t; and (c) holds since conditioning reduces entropy.

Therefore, we have

$$\begin{aligned}
&\sum_{t=1}^{n} H(G_{12}[t]X_1[t]|W_2, G_{12}^{t-1}X_1^{t-1}, X_2^t, G^t)\\
&\quad\geq \frac{1}{2-p}\sum_{t=1}^{n} H(G_{11}[t]X_1[t], G_{12}[t]X_1[t]|W_2, G_{11}^{t-1}X_1^{t-1}, G_{12}^{t-1}X_1^{t-1}, X_2^t, G^t).
\end{aligned}$$

■

In the following two sections, we consider the last scenario we are interested in, *i.e.* Instantaneous-CSIT and OFB, and we provide the proof of Theorem 4. First, we present the achievability strategy, and we demonstrate how OFB can enhance our achievable rate region. We then present the converse proof.

IX. Achievability Proof of Theorem 4 [Instantaneous-CSIT and OFB]

In this section, we describe our achievability strategy for the case of Instantaneous-CSIT and output feedback. Note that in this scenario, although transmitters have instantaneous knowledge of the channel state information, the output signals are available at the transmitters with unit delay. We first provide a brief overview of our scheme.

A. Overview

By symmetry, it suffices to describe the achievability scheme for corner point

$$(R_1, R_2) = \left(1-q^2, pq\right),$$

as depicted in Figure 26. Similarly, we can achieve corner point $(R_1, R_2) = \left(pq, 1-q^2\right)$, and therefore by time sharing, we can achieve the region.

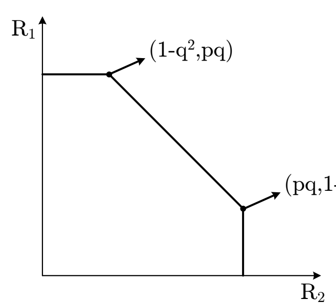


Fig. 26. Two-user Binary Fading IC: capacity region with Instantaneous-CSIT and output feedback. By symmetry, it suffices to describe the achievability scheme for corner point $(R_1, R_2) = \left(1-q^2, pq\right)$.

Our achievability strategy is carried on over $b+1$ communication blocks, each block with n time instants. Transmitters communicate fresh data bits in the first b blocks and the final block is to help the receivers decode their corresponding bits. At the end, using our scheme, we achieve a rate tuple arbitrary close to $\frac{b}{b+1}\left(1-q^2, pq\right)$ as $n \to \infty$. Finally letting $b \to \infty$, we achieve the desired rate tuple.

B. Achievability Strategy

Let W_i^j be the message of transmitter i in block j, $i = 1, 2$, and $j = 1, 2, \ldots, b$. Moreover, let $\mathrm{W}_1^j = a_1^j, a_2^j, \ldots, a_m^j$, and $\mathrm{W}_2^j = b_1^j, b_2^j, \ldots, b_{m_2}^j$, for $j = 1, 2, \ldots, b$, where data bits a_i^j's and b_i^j's are picked uniformly and independently from $\{0, 1\}$, $i = 1, 2, \ldots, m$, and

$$m_2 = \frac{q}{1+q} m. \tag{62}$$

We also set $n = m/(1-q^2) + m^{2/3}$, where n is the length of each block.

Achievability strategy for block 1: In the first communication block, at each time instant t, if at least one of the outgoing links from Tx_1 is on, then it sends one of its initial m bits that has not been transmitted before (note that this happens with probability $(1-q^2)$). On the other hand, Tx_2 communicates a new bit (a bit that has not been transmitted before) if the link to its receiver is on and it does not interfere with receiver one (*i.e.* $G_{22}[t] = 1$ and $G_{21}[t] = 0$). In other words, Tx_2 communicates a new bit if either one of Cases $2, 4, 9$, or 11 in Table II occurs (note that this happens with probability pq).

The first block goes on for n time instants. If at the end of the first block, there exists a bit at either of the transmitters that has not yet been transmitted, we consider it as error type-I and halt the transmission.

Assuming that the transmission in not halted, using output feedback links, transmitter two has access to the bits of transmitter one communicated in the first block. In particular, Tx_2 has access to the bits of Tx_1 transmitted in Cases $2, 11, 12, 14$, and 15 during block 1. Note that the bits communicated in Cases $11, 12, 14$, and 15 from Tx_1 have to be provided to Rx_1. However, the bits communicated in Case 2 from Tx_1 are already available at Rx_1 but needed at Rx_2, see Figure 27. Transmitter two will provide such bits to Rx_2 in the following communication block.

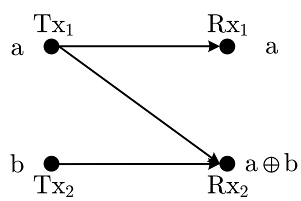


Fig. 27. The bit communicated in Case 2 from Tx_1 is already available at Rx_1 but it is needed at Rx_2. Transmitter two learns this bit through the feedback channel and will provide it to Rx_2 in the following communication block.

Now, Tx_2 transfers the bits of Tx_1 communicated in Cases $2, 11, 12, 14$, and 15, during the first communication block to queues Q^1_{1,C_2}, $Q^1_{1,\mathsf{C}_{11}}$, $Q^1_{1,\mathsf{C}_{12}}$, $Q^1_{1,\mathsf{C}_{14}}$, and $Q^1_{1,\mathsf{C}_{15}}$ respectively.

Let random variable N^1_{1,C_ℓ} denote the number of bits in Q^1_{1,C_ℓ}, $\ell = 2, 11, 12, 14, 15$. Since transition of a bit to this state is distributed as independent Bernoulli RV, upon completion of block 1, we have

$$\mathbb{E}[N^1_{1,\mathsf{C}_\ell}] = \frac{\Pr\left(\text{Case } \ell\right)}{1 - \sum_{i=9,10,13,16} \Pr\left(\text{Case i}\right)} m = (1-q^2)^{-1} \Pr\left(\text{Case } \ell\right) m, \qquad \ell = 2, 11, 12, 14, 15. \tag{63}$$

If the event $\left[N^1_{1,\mathsf{C}_\ell} \geq \mathbb{E}[N^1_{1,\mathsf{C}_\ell}] + m^{\frac{2}{3}}\right]$ occurs, we consider it as error type-II and we halt the transmission. At the end of block 1, we add 0's (if necessary) to Q^1_{1,C_ℓ} so that the total number of bits is equal to $\mathbb{E}[N^1_{1,\mathsf{C}_\ell}] + m^{\frac{2}{3}}$. Furthermore, using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II decreases exponentially with m.

Achievability strategy for block j, $j = 2, 3, \ldots, b$: The transmission strategy for Tx_1 is the same as block 1 for the first b blocks (all but the last block). In other words, at time instant t, Tx_1 transmits one of its initial m bits (that has not been transmitted before) if at least one of its outgoing links is on. On the other hand, Tx_2 communicates W_2^2 using similar strategy as the first block, *i.e.* Tx_2 communicates a new bit if either one of Cases $2, 4, 9$, or 11 occurs.

Transmitter two transfers the bits communicated in Cases $2, 11, 12, 14$, and 15, during communication block j to queues Q^j_{1,C_2}, $Q^j_{1,\mathsf{C}_{11}}$, $Q^j_{1,\mathsf{C}_{12}}$, $Q^j_{1,\mathsf{C}_{14}}$, and $Q^j_{1,\mathsf{C}_{15}}$ respectively.

Moreover, at time instant t,

- if Case 3 occurs, Tx_2 sends one of the bits from Q^{j-1}_{1,C_2} and removes it from this queue since it has been delivered successfully to Rx_2, see Figure 28. If Case 3 occurs and Q^{j-1}_{1,C_2} is empty, Tx_2 remains silent;

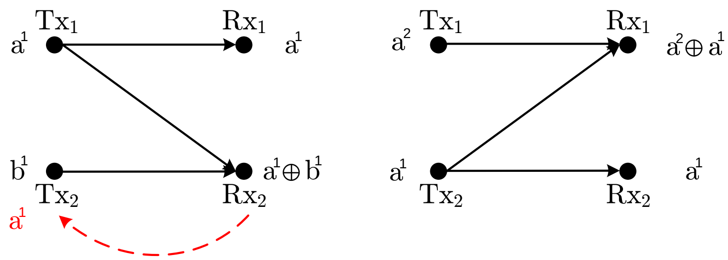


Fig. 28. In block j when Case 3 occurs, Tx_2 retransmits the bit of Tx_1 communicated in Case 2 during block $j-1$. Note that this bit does not cause interference at Rx_1 and it is needed at Rx_2.

- if Case 10 occurs, Tx_2 sends one of the bits from $Q^{j-1}_{1,\mathsf{C}_{11}}$ and removes it from this queue, see Figure 29. If Case 10 occurs and $Q^{j-1}_{1,\mathsf{C}_{11}}$ is empty, Tx_2 remains silent;

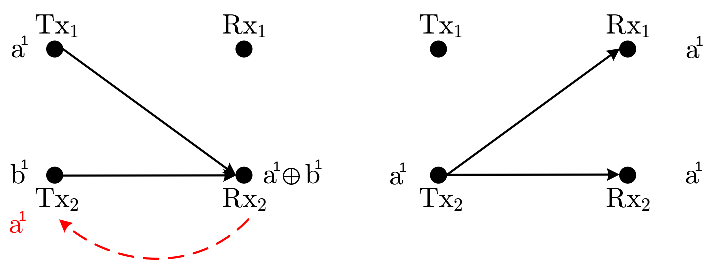


Fig. 29. In block j when Case 10 occurs, Tx_2 retransmits the bit of Tx_1 communicated in Case 11 during block $j-1$. Note that this bit is needed at both receivers.

- if Case 12 occurs, it sends one of the bits from $Q^{j-1}_{1,\mathsf{C}_{12}}$ and removes it from this queue. If Case 12 occurs and $Q^{j-1}_{1,\mathsf{C}_{12}}$ is empty, Tx_2 remains silent;
- if Case 13 occurs, it sends one of the bits from $Q^{j-1}_{1,\mathsf{C}_{14}}$ and removes it from this queue. If Case 13 occurs and $Q^{j-1}_{1,\mathsf{C}_{14}}$ is empty, Tx_2 remains silent;
- if Case 15 occurs, it sends one of the bits from $Q^{j-1}_{1,\mathsf{C}_{15}}$ and removes it from this queue. If Case 15 occurs and $Q^{j-1}_{1,\mathsf{C}_{15}}$ is empty, Tx_2 remains silent.

If at the end of block j, there exists a bit at either of the transmitters that has not yet been transmitted, or any of the queues Q^{j-1}_{1,C_2}, $Q^{j-1}_{1,\mathsf{C}_{11}}$, $Q^{j-1}_{1,\mathsf{C}_{12}}$, $Q^{j-1}_{1,\mathsf{C}_{14}}$, or $Q^{j-1}_{1,\mathsf{C}_{15}}$ is not empty, we consider this event as error type-I and halt the transmission.

Assuming that the transmission is not halted, let random variable N^j_{1,C_ℓ} denote the number of bits in Q^j_{1,C_ℓ}, $\ell = 2, 11, 12, 14, 15$. From basic probability, we have

$$\mathbb{E}[N^j_{1,\mathsf{C}_\ell}] = \frac{\Pr\left(\text{Case } \ell\right)}{1 - \sum_{i=9,10,13,16} \Pr\left(\text{Case i}\right)} m = (1-q^2)^{-1} \Pr\left(\text{Case } \ell\right) m, \qquad \ell = 2, 11, 12, 14, 15. \tag{64}$$

If the event $\left[N^j_{1,\mathsf{C}_\ell} \geq \mathbb{E}[N^j_{1,\mathsf{C}_\ell}] + m^{\frac{2}{3}} \right]$ occurs, we consider it as error type-II and we halt the transmission. At the end of block 1, we add 0's (if necessary) to Q^j_{1,C_ℓ} so that the total number of bits is equal to $\mathbb{E}[N^j_{1,\mathsf{C}_\ell}] + m^{\frac{2}{3}}$. Using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II and decreases exponentially with m.

Achievability strategy for block $b+1$: Finally in block $b+1$, no new data bit is transmitted (*i.e.* $\mathrm{W}_1^{b+1}, \mathrm{W}_2^{b+1} = 0$), and Tx_2 only communicates the bits of Tx_1 communicated in the previous block in Cases $2, 11, 12, 14$, and 15 as

described above. If at the end of block $b+1$, any of the queues Q^b_{1,C_2}, $Q^b_{1,\mathsf{C}_{11}}$, $Q^b_{1,\mathsf{C}_{12}}$, $Q^b_{1,\mathsf{C}_{14}}$, or $Q^b_{1,\mathsf{C}_{15}}$ is not empty, we consider this event as error type-I and halt the transmission.

The probability that the transmission strategy halts at the end of each block can be bounded by the summation of error probabilities of types I and II. Using Chernoff-Hoeffding bound, we can show that the probability that the transmission strategy halts at any point approaches zero as $m \rightarrow \infty$.

C. Decoding

At the end of block $j+1$, Rx_1 has acces to W_1^j with no interference, $j = 1, 2, \ldots, b$. At the end of block $b+1$, Rx_2 uses the bits communicated in Cases 3 and 10 from Tx_2 to cancel out the interference it has received from Tx_1 during the previous block in Cases 2 and 11. Therefore, at the end of block $b+1$, Rx_1 has access to W_2^b with no interference. Then, Rx_2 follows the same strategy for blocks b and $b-1$. Therefore, using similar idea, Rx_2 uses backward decoding to cancel out interference in the previous blocks to decode all messages.

Now, since each block has $n = m/(1-q^2)+m^{2/3}$ time instants and the probability that the transmission strategy halts at any point approaches zero for $m \rightarrow \infty$, we achieve a rate tuple

$$\frac{b}{b+1}\left(1-q^2, pq\right), \tag{65}$$

as $m \rightarrow \infty$. Finally letting $b \rightarrow \infty$, we achieve the desired rate tuple.

X. Converse Proof of Theorem 4 [Instantaneous-CSIT and OFB]

To derive the outer-bound on individual rates, we have

$$\begin{aligned} nR_1 &= H(W_1) \overset{(a)}{=} H(W_1|G^n) \\ &\overset{\text{(Fano)}}{\leq} I(W_1; Y_1^n|G^n) + n\epsilon_n \\ &= H(Y_1^n|G^n) - H(Y_1^n|W_1, G^n) + n\epsilon_n \\ &\leq H(Y_1^n|G^n) + n\epsilon_n \\ &\leq (1-q^2)n + n\epsilon_n, \end{aligned} \tag{66}$$

where $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$; (a) holds since message W_1 is independent of G^n. Similarly, we have

$$nR_2 \leq (1-q^2)n + n\epsilon_n, \tag{67}$$

dividing both sides by n and let $n \rightarrow \infty$, we have

$$\begin{cases} R_1 \leq 1 - q^2, \\ R_2 \leq 1 - q^2. \end{cases} \tag{68}$$

The outer-bound on $R_1 + R_2$, *i.e.*

$$R_1 + R_2 \leq 1 - q^2 + pq, \tag{69}$$

can be obtained as follows.

$$
\begin{aligned}
& n(R_1+R_2-2\epsilon_n) \\
& \stackrel{(a)}{\leq} H(W_1|W_2,G^n)+H(W_2|G^n) \\
& \stackrel{\text{Fano}}{\leq} I(W_1;Y_1^n|W_2,G^n)+I(W_2;Y_2^n|G^n) \\
& = H(Y_1^n|W_2,G^n) - \underbrace{H(Y_1^n|W_1,W_2,G^n)}_{=\,0} + I(W_2;Y_2^n|G^n) \\
& = H(Y_1^n|W_2,G^n)+H(Y_2^n|G^n)-H(Y_2^n|W_2,G^n) \\
& = H(Y_1^n|W_2,G^n)+H(Y_2^n|G^n)-\left[H(Y_1^n,Y_2^n|W_2,G^n)-H(Y_1^n|Y_2^n,W_2,G^n)\right] \\
& = H(Y_1^n|Y_2^n,W_2,G^n)+H(Y_2^n|G^n) \\
& \stackrel{(b)}{=} H(Y_2^n|G^n)+\sum_{t=1}^{n} H(Y_1[t]|W_2,Y_2^n,Y_1^{t-1},X_2^t,G_{12}^tX_1^t,G^n) \\
& \stackrel{(c)}{\leq} H(Y_2^n|G^n)+H(Y_1^n|G_{12}^nX_1^n,G_{21}^nX_2^n,G^n) \\
& \stackrel{(d)}{\leq} \sum_{t=1}^{n} H(Y_2[t]|G^n)+\sum_{t=1}^{n} H(Y_1[t]|G_{12}[t]X_1[t],G_{21}[t]X_2[t],G^n) \\
& \stackrel{(e)}{\leq} \left(1-q^2\right)n+pqn,
\end{aligned} \tag{70}
$$

where $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$; and (a) follows from the fact that the messages and G^n are mutually independent; (b) holds since X_2^t is a function of $W_2, Y_1^{t-1}, Y_2^{t-1}$, and G^t; (c) and (d) follow from the fact that conditioning reduces entropy; and (e) holds since

$$
\begin{aligned}
& H(Y_2[t]|G^n) \leq 1-q^2, \\
& H(Y_1[t]|G_{12}[t]X_1[t],G_{21}[t]X_2[t],G^n) \leq pq.
\end{aligned} \tag{71}
$$

Dividing both sides by n and let $n \rightarrow \infty$, we get

$$
R_1+R_2 \leq 1-q^2+pq. \tag{72}
$$

We also note this outer-bound on R_1+R_2 can be also applied to the case of Instantaneous-CSIT and no output feedback (*i.e.* Theorem 1).

XI. Conclusion and Future Directions

We studied the impact of delayed knowledge of the channel state information at the transmitters, on the capacity region of the two-user binary fading interference channels. We introduced various coding opportunities, created by Delayed-CSIT, and presented an achievability strategy that systematically exploits the coding opportunities. We derived an achievable rate region that matches the outer-bounds for this problem, hence, characterizing the capacity region. We have also derived the capacity region of this problem with Delayed-CSIT and output feedback.

A future direction would be to extend our results to the case of two-user Gaussian fading interference channel with Delayed-CSIT. As discussed in the introduction, one can view our binary fading model as a fading interpretation of the linear deterministic model where the non-negative integer associated to each link is at most 1. Therefore, one approach is to extend the current results to the case of fading linear deterministic interference channel and then, further extend that result to the case of Gaussian fading interference channel, in order to obtain approximate capacity characterization.

Another future direction is to consider the k-user setting of the problem. In [23], authors have shown that for the k-user fading interference channel with instantaneous knowledge of the channel state information, sum degrees of freedom (DoF) of $k/2$ is achievable. However, in the absence of the CSIT, the achievable sum DoF collapses to 1. As a result a large degradation in network capacity, due to lack of the CSIT, is observed. It has been recently shown that, with Delayed-CSIT, it is possible to achieve more than one sum DoF [7], [8], however, the achievable

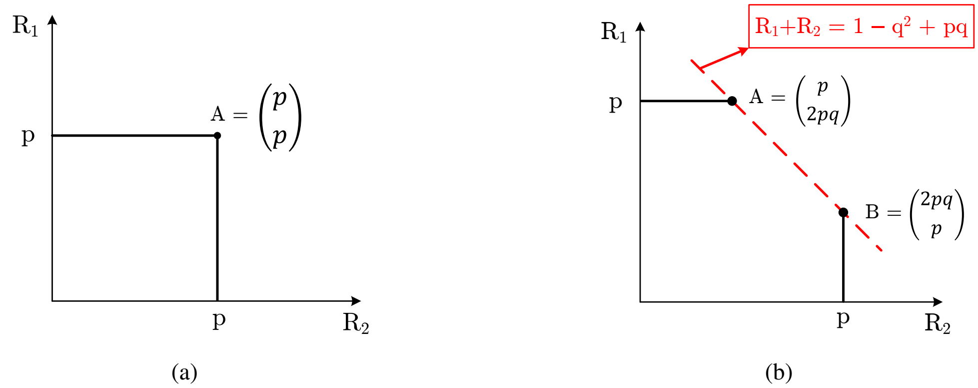


Fig. 30. *Capacity region of the two-user BFIC with Instantaneous-CSIT and for (a)* $0 \leq p \leq 0.5$*, and (b)* $0.5 < p \leq 1$.

sum DoFs are less than 1.5 for any number of users. This together with lack of nontrivial DoF upper bounds leaves the problem of sum DoF characterization of interference channels with Delayed-CSIT still open and challenging, to the extent that it is even unknown whether the sum DoF of such networks *scales* with the number of users or not. A promising direction may be to study this problem in the context of our simpler binary fading model, to understand whether the sum capacity of such network with Delayed-CSIT scales with the number of users or it will saturate.

Finally, motivated by recent results that demonstrate that, with Instantaneous-CSIT, multi-hopping can significantly increase the capacity of interference networks (*e.g.*, [24], [25] for two-unicast networks and [26] for multi-unicast networks), an interesting future direction would be explore the impact of Delayed-CSIT on the capacity of muti-hop binary interference networks.

Appendix A
Achievability Proof of Theorem 1 [Instantaneous-CSIT]

In this appendix, we provide the achievability proof of Theorem 1. Below, we have stated the capacity region of the two-user BFIC with Instantaneous-CSIT (and no OFB).

$$\mathcal{C}^{\text{ICSIT}} = \begin{cases} 0 \leq R_i \leq p, & i = 1, 2, \\ R_1 + R_2 \leq 1 - q^2 + pq. & \end{cases} \tag{73}$$

Remark 14: For $0 \leq p \leq 0.5$, the capacity region is given by

$$\mathcal{C}^{\text{ICSIT}} = \left\{ R_1, R_2 \in \mathbb{R}^+ \ s.t. \ R_i \leq p, \ i = 1, 2 \right\}. \tag{74}$$

while for $0.5 < p \leq 1$, the outer-bound on $R_1 + R_2$ is also active, see Figure 30.

With Instantaneous-CSIT, each transmitter knows what channel realization occurs at the time of transmission. Transmitters can take advantage of such knowledge and by pairing different realizations, the optimal rate region as given in Theorem 1 can be achieved. We will first describe the achievability strategy for $0 \leq p \leq 0.5$, since it is easier to follow. We then complete the proof by describing the achievability strategy for $0.5 < p \leq 1$.

A. Achievabiliy Strategy for $0 \leq p \leq 0.5$

Note that the result for $p = 0$ is trivial, so we assume $0 < p \leq 0.5$. Below, we describe the possible pairing opportunities that are useful in this regime and then, we describe the achievability scheme. The possible pairing opportunities are as follows.

- Type A [Cases 1 and 15]: In Case 15, only the corss links are equal to 1, therefore, by pairing bits in Case 1 with bits in Case 15, we can cancel out interference in Case 1, see Figure 31. In other words by pairing the two cases, we can communicate 2 bits interference free.
- Type B [Cases 2 and 14]: We can pair up Cases 2 and 14 to cancel out interference in Case 2 as depicted in Figure 32.

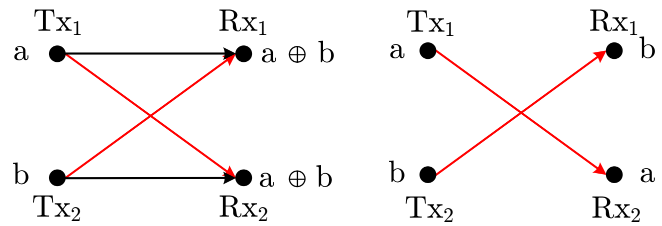


Fig. 31. Pairing opportunity Type A: By pairing Cases 1 and 15, we can communicate two bits interference-free. For instance, receiver one has access to bits $a \oplus b$ and b and as a result, it can decode its desired bit.

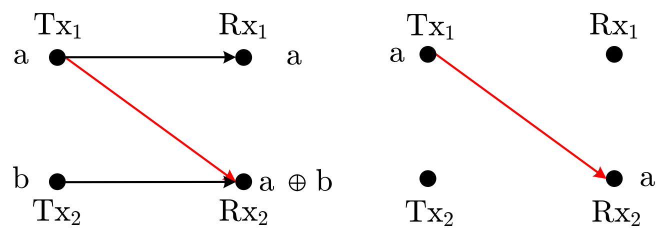


Fig. 32. Pairing opportunity Type B: By pairing Cases 2 and 14, we can communicate two bits interference-free. For instance, receiver two has access to bits $a \oplus b$ and a and as a result, it can decode its desired bit.

- Type C [Cases 3 and 13]: Similar to Type B with swapping user IDs.

We are now ready to provide the achievability scheme for the Instantaneous-CSIT model and for $0 \leq p \leq 0.5$. We first provide an overview of our scheme.

1) Overview: Our achievability strategy is carried on over $b+1$ communication blocks, each block with n time instants. We describe the achievability strategy for rate tuple

$$(R_1, R_2) = (p, p). \tag{75}$$

Transmitters communicate fresh data bits in the first b blocks and the final block is to help the receivers decode their corresponding bits. At the end, using our scheme, we achieve rate tuple $\frac{b}{b+1}(p, p)$ as $n \to \infty$. Finally, letting $b \to \infty$, we achieve the desired rate tuple. In our scheme the messages transmitted in block j, $j = 1, 2, \ldots, b$, will be decoded at the end of block $j+1$.

2) Achievability strategy: Let W_i^j be the message of transmitter i in block j, $i = 1, 2$, $j = 1, 2, \ldots, b$. Moreover, let $\mathrm{W}_1^j = a_1^j, a_2^j, \ldots, a_m^j$, and $\mathrm{W}_2^j = b_1^j, b_2^j, \ldots, b_m^j$, where a_i^j's and b_i^j's are picked uniformly and independently from $\{0, 1\}$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, b$, and some positive integer m. We set

$$n = m/p + \left(2/p^4\right) m^{2/3}. \tag{76}$$

Achievability strategy for block 1: In the first communication block, at each time instant t, Tx_i sends a new data bit (from its initial m bits) if $G_{ii}[t] = 1$, $i = 1, 2$. In other words, Tx_1 sends a new data bit either of the following channel realizations occurs (see Table II): Cases $1, 2, 3, 4, 5, 6, 7$, and 8; while Tx_2 sends a new data bit if either of the following channel realizations occurs: Cases $1, 2, 3, 4, 9, 10, 11$, and 12.

If not specified, the transmitters remain silent. Tx_1 transfers its transmitted bits in Cases 1 and 2 to queues $Q_{1,\mathsf{C1}}^1$ and $Q_{1,\mathsf{C2}}^1$ respectively; and Tx_2 transfers its transmitted bits in Cases 1 and 3 to queues $Q_{2,\mathsf{C1}}^1$ and $Q_{2,\mathsf{C3}}^1$ respectively.

If at the end of block 1, there exists a bit at either of the transmitters that has not yet been transmitted, we consider it as error type-I and halt the transmission.

Remark 15: Note that the transmitted bits in Cases $4, 5, 6, 7, 8, 9, 10, 11$, and 12 are available at their corresponding receivers without any interference. In other words, they are communicated successfully and no retransmission is required.

Assuming that the transmission is not halted, let random variable N^1_{i,C_ℓ} denote the number of bits in Q^1_{i,C_ℓ}, $(i,\ell) = (1,1),(1,2),(2,1),(2,3)$. Since transition of a bit to this queue is distributed as independent Bernoulli RV, upon completion of block 1, we have

$$\mathbb{E}[N^1_{i,\mathsf{C}_\ell}] = \frac{\Pr\left(\text{Case } \ell\right)}{1 - \sum_{i=9,10,\ldots,16} \Pr\left(\text{Case i}\right)} m = \frac{1}{p} \Pr\left(\text{Case } \ell\right) m. \tag{77}$$

If the event $\left[N^1_{i,\mathsf{C}_\ell} \geq \mathbb{E}[N^1_{i,\mathsf{C}_\ell}] + m^{\frac{2}{3}}\right]$ occurs, we consider it as error type-II and we halt the transmission. At the end of block 1, we add 0's (if necessary) to Q^1_{i,C_ℓ} so that the total number of bits is equal to $\mathbb{E}[N^1_{i,\mathsf{C}_\ell}] + m^{\frac{2}{3}}$. Furthermore, using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II decreases exponentially with m.

Achievability strategy for block j, $j = 2, 3, \ldots, b$: In communication block j, $j = 2, 3, \ldots, b$, at each time instant t, Tx_i sends a new data bit (from its initial m bits) if $G_{ii}[t] = 1$, $i = 1, 2$. Transmitter one transfers its transmitted bit in Cases 1 and 2 to queues $Q^j_{1,\mathsf{C1}}$ and $Q^j_{1,\mathsf{C2}}$ respectively; and Tx_2 transfers its transmitted bit in Cases 1 and 3 to queues $Q^j_{2,\mathsf{C1}}$ and $Q^j_{2,\mathsf{C3}}$ respectively. Note that so far the transmission scheme is similar to the first communication block.

Now if at a given time instant Case 15 occurs, Tx_i sends a bit from queue $Q^{j-1}_{i,\mathsf{C1}}$ and removes it from the this queue. If at time instant t Case 15 occurs and $Q^{j-1}_{i,\mathsf{C1}}$ is empty, then Tx_i remains silent. This way, similar to pairing Type A described previously, the transmitted bits in Case 1 of the previous block can be decoded at the corresponding receiver.

Furthermore, if at a given time instant Case 14 (13) occurs, Tx_1 (Tx_2) sends a bit from queue $Q^{j-1}_{1,\mathsf{C2}}$ $(Q^{j-1}_{2,\mathsf{C3}})$ and removes it from the this queue. This is motivated by pairing Type B (C) described previously.

If at the end of block j, there exists a bit at either of the transmitters that has not yet been transmitted, or any of the queues Q^{j-1}_{1,C_1}, Q^{j-1}_{1,C_2}, Q^{j-1}_{2,C_1}, or Q^{j-1}_{2,C_3} is not empty, we consider this event as error type-I and we halt the transmission.

Assuming that the transmission is not halted, let random variable N^j_{i,C_ℓ} denote the number of bits in Q^j_{i,C_ℓ}, $(i,\ell) = (1,1),(1,2),(2,1),(2,3)$. Since transition of a bit to this state is distributed as independent Bernoulli RV, upon completion of block j, we have

$$\mathbb{E}[N^j_{i,\mathsf{C}_\ell}] = \frac{\Pr\left(\text{Case } \ell\right)}{1 - \sum_{i=9,10,\ldots,16} \Pr\left(\text{Case i}\right)} m = \frac{1}{p} \Pr\left(\text{Case } \ell\right) m. \tag{78}$$

If the event $\left[N^j_{i,\mathsf{C}_\ell} \geq \mathbb{E}[N^j_{i,\mathsf{C}_\ell}] + m^{\frac{2}{3}}\right]$ occurs, we consider it as error type-II and we halt the transmission. At the end of block 1, we add 0's (if necessary) to Q^j_{i,C_ℓ} so that the total number of bits is equal to $\mathbb{E}[N^j_{i,\mathsf{C}_\ell}] + m^{\frac{2}{3}}$. Using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II decreases exponentially with m.

Achievability strategy for block $b+1$: In the final communication block, transmitters do not communicate any new data bit.

If at time instant t Case 15 occurs, Tx_i sends a bit from queue $Q^b_{i,\mathsf{C1}}$ and removes it from the this queue. If at time instant t Case 15 occurs and $Q^b_{i,\mathsf{C1}}$ is empty, then Tx_i remains silent. If at time instant t Case 14 (13) occurs, Tx_1 (Tx_2) sends a bit from queue $Q^b_{1,\mathsf{C2}}$ $(Q^b_{2,\mathsf{C3}})$ and removes it from the this queue.

If at the end of block $b+1$, any of the states Q^b_{1,C_1}, Q^b_{1,C_2}, Q^b_{2,C_1}, or Q^b_{2,C_3} is not empty, we consider this event as error type-I and we halt the transmission.

Note that if the transmission is not halted, any bit is either available at its intended receiver interference-free, or the interfering bit is provided to the receiver in the following block. The probability that the transmission strategy halts at the end of each block can be bounded by the summation of error probabilities of types I and II. Using Chernoff-Hoeffding bound, we can show that the probability that the transmission strategy halts at any point approaches zero as $m \to \infty$.

Now, since each block has $n = m/p + \left(2/p^4\right) m^{2/3}$ time instants and the probability that the transmission strategy halts at any point approaches zero as $m \to \infty$, we achieve a rate tuple

$$\frac{b}{b+1}\left(p, p\right), \tag{79}$$

as $m \to \infty$. Finally letting $b \to \infty$, we achieve the desired rate tuple.

B. Achievabiliy Strategy for $0.5 < p \leq 1$

For $p = 1$, the capacity region is the same with no, delayed, or instantaneous CSIT. So in this section, we assume $0.5 < p < 1$. By symmetry, it suffices to describe the strategy for point $A = (p, 2pq)$. In this regime, we will take advantage of another pairing opportunity as described below.

• Type D [Cases $2, 3$, and 12]: This type of pairing is different from what we have described so far. In all previous types, we paired up cases that had zero capacity to cancel out interference in other cases. However, here all three cases have capacity 1. By pairing all three cases together, we can communicate 4 bits as depicted in Figure 33.

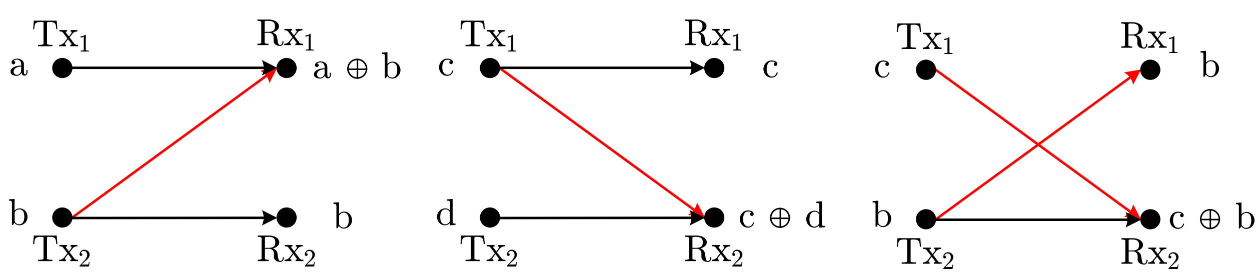


Fig. 33. Pairing opportunity Type D: Cases $2, 3$, and 12. Tx_1 uses c to recover b and then it decodes a, similar argument holds for Tx_2. All three cases have capacity 1, and by pairing them, we can communicate 4 bits.

1) Overview: The achievability is again carried on over $b+1$ communication blocks, each block with n time instants. We describe the achievability strategy for rate tuple

$$\left(R_1, R_2\right) = \left(p, 2pq\right), \tag{80}$$

see Figure 30(b).

Transmitters communicate fresh data bits in the first b blocks and the final block is to help receivers decode their corresponding bits. At the end, using our scheme, we achieve rate tuple $\frac{b}{b+1}\left(p, 2pq\right)$ as $n \to \infty$. Finally, letting $b \to \infty$, we achieve the desired corner point. In our scheme, the transmitted bits in block j, $j = 1, 2, \ldots, b$, will be decoded by the end of block $j+1$.

2) Achievability strategy: Let W_i^j be the message of transmitter i in block j. We assume $\mathrm{W}_1^j = a_1^j, a_2^j, \ldots, a_m^j$, and $\mathrm{W}_2^j = b_1^j, b_2^j, \ldots, b_{m_2}^j$ for $j = 1, 2, \ldots, b$, where a_i^j's and b_i^j's are picked uniformly and independently from $\{0, 1\}$, for some positive value of m and $m_2 = 2qm$ (note that $2q < 1$). We set

$$n = m/p + \left(2/q^4\right) m^{2/3}. \tag{81}$$

Achievability strategy for block 1: In the first communication block, at each time instant t, transmitter one sends a new data bit if $G_{11}[t] = 1$ except Case 1. In other words, Tx_1 sends a new data bit if either of the following channel realizations occurs (see Table II): Cases $2, 3, 4, 5, 6, 7$, and 8. Transmitter two sends a new data bit if $G_{22}[t] = 1$ except Cases 1 and 12. In other words, Tx_2 sends a new data bit if either of the following channel realizations occurs: Cases $2, 3, 4, 9, 10$, and 11.

If at time instant t where $t \leq \frac{q^2}{p^2} n$, Case 1 occurs, then each transmitter sends out a new data bit. Then, Tx_i transfers its transmitted bit in Case 1 to queue $Q_{i,\mathsf{C}1}^1$ for $t \leq \frac{q^2}{p^2} n$. If $t > \frac{q^2}{p^2} n$ and Case 1 occurs, then Tx_1 sends out a new data bit while Tx_2 remains silent, see Figure 34. Note that these bits are delivered to Rx_1 interference-free.

If $t \leq \frac{q^2}{p^2} n$, and Case 12 occurs, then Tx_2 sends out a new data bit while Tx_1 remains silent. Note that these bits are delivered to Rx_2 interference-free.

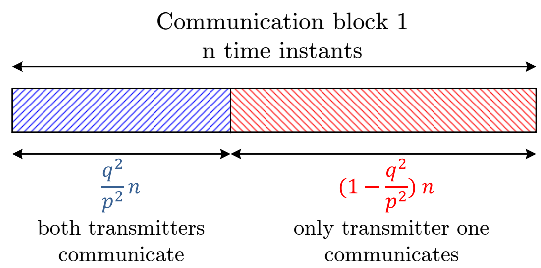


Fig. 34. If Case 1 occurs during communication block 1, then if $t \leq \frac{q^2}{p^2}n$, each transmitter sends out a new data bit. However, if $t > \frac{q^2}{p^2}n$, then Tx_1 sends out a new data bit while Tx_2 remains silent.

If not specified, the transmitters remain silent. Note that Tx_1 sends a bit if $G_{11}[t] = 1$ (*i.e.* with probability p). On the other hand, Tx_2 sends a bit with probability

$$\sum_{j=2,3,4,9,10,11} \Pr\left(\text{Case j}\right) + \frac{q^2}{p^2} \sum_{j=1,12} \Pr\left(\text{Case j}\right) = 2pq. \tag{82}$$

Transmitter one transfers its transmitted bit in Case 2 to queue $Q^1_{1,\mathsf{C2}}$; and Tx_2 transfers its transmitted bit in Case 3 to queue $Q^1_{2,\mathsf{C3}}$. If at the end of block 1, there exists a bit at either of the transmitters that has not yet been transmitted, we consider it as error type-I and halt the transmission.

Remark 16: Note that the transmitted bits in Cases $4, 5, 6, 7, 8, 9, 10$, and 11 are available at their corresponding receivers without any interference.

Assuming that the transmission is not halted, let random variable N^1_{i,C_ℓ} denote the number of bits in Q^1_{i,C_ℓ}, $(i,\ell) = (1,1), (1,2), (2,1), (2,3)$. Since transition of a bit to this state is distributed as independent Bernoulli RV, upon completion of block 1, we have

$$\begin{aligned}
\mathbb{E}[N^1_{1,\mathsf{C}_1}] &= \frac{\left(q^2/p^2\right)\Pr\left(\text{Case 1}\right)}{1-\sum_{j=9,10,\ldots,16}\Pr\left(\text{Case j}\right)}m = pq^2 m,\\
\mathbb{E}[N^1_{1,\mathsf{C}_2}] &= \frac{\Pr\left(\text{Case 2}\right)}{1-\sum_{j=9,10,\ldots,16}\Pr\left(\text{Case j}\right)}m = p^2 q m,\\
\mathbb{E}[N^1_{2,\mathsf{C}_1}] &= \frac{\left(q^2/p^2\right)\Pr\left(\text{Case 1}\right)}{\sum_{j=2,3,4,9,10,11}\Pr\left(\text{Case j}\right)+\frac{q^2}{p^2}\sum_{j=1,12}\Pr\left(\text{Case j}\right)}2qm = pq^2 m,\\
\mathbb{E}[N^1_{2,\mathsf{C}_3}] &= \frac{\Pr\left(\text{Case 3}\right)}{\sum_{j=2,3,4,9,10,11}\Pr\left(\text{Case j}\right)+\frac{q^2}{p^2}\sum_{j=1,12}\Pr\left(\text{Case j}\right)}2qm = p^2 q m.
\end{aligned} \tag{83}$$

If the event $\left[N^1_{i,\mathsf{C}_\ell} \geq \mathbb{E}[N^1_{i,\mathsf{C}_\ell}] + m^{\frac{2}{3}}\right]$ occurs, we consider it as error type-II and we halt the transmission. At the end of block 1, we add 0's (if necessary) to Q^1_{i,C_ℓ} so that the total number of bits is equal to $\mathbb{E}[N^1_{i,\mathsf{C}_\ell}] + m^{\frac{2}{3}}$. Using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II decreases exponentially with m.

Achievability strategy for block j, $j = 2, 3, \ldots, b$: In communication block j, $j = 2, 3, \ldots, b$, at each time instant t, transmitter one sends a new data bit if $G_{11}[t] = 1$ except Case 1, while transmitter two sends a new data bit if $G_{22}[t] = 1$ except Cases 1 and 12.

If $t \leq \frac{q^2}{p^2}n$ and Case 1 occurs, then each transmitter sends out a new data bit. Then Tx_i transfers its transmitted bit in Case 1 to queue $Q^j_{i,\mathsf{C1}}$ for $t \leq \frac{q^2}{p^2}n$. If $t > \frac{q^2}{p^2}n$ and Case 1 occurs, then Tx_1 sends out a new data bit while Tx_2 remains silent. Note that these bits are delivered to Rx_1 interference-free.

If $t \leq \frac{q^2}{p^2}n$ and Case 12 occurs, then Tx_2 sends out a new data bit while Tx_1 remains silent. We will exploit channel realization 12 for $t > \frac{q^2}{p^2}n$, to perform pairing Type D.

Transmitter one transfers its transmitted bit in Case 2 to queue $Q^j_{1,\mathsf{C2}}$; and transmitter two transfers its transmitted bit in Case 3 to queue $Q^j_{2,\mathsf{C3}}$. Note that so far the transmission scheme is similar to the first communication block.

Now, if at time instant t Case 15 occurs, Tx_i sends a bit from queue $Q_{i,\mathsf{C1}}^{j-1}$ and removes it from the this queue. If at time instant t Case 15 occurs and $Q_{i,\mathsf{C1}}^{j-1}$ is empty, then Tx_i remains silent. This way, similar to pairing Type A described previously, the transmitted bits in Case 1 of the previous block can be decoded at the corresponding receiver.

Furthermore, if at time instant t Case 14 (13) occurs, Tx_1 (Tx_2) sends a bit from queue $Q_{1,\mathsf{C2}}^{j-1}$ ($Q_{2,\mathsf{C3}}^{j-1}$) and removes it from the this queue. This is motivated by pairing Type B (C) described previously.

Finally, if $t > \frac{q^2}{p^2}n$ and Case 12 occurs, Tx_1 sends a bit from queue $Q_{1,\mathsf{C2}}^{j-1}$ and Tx_2 sends a bit from queue $Q_{2,\mathsf{C3}}^{j-1}$. Each transmitter removes the transmitted bit from the corresponding queue. This is motivated by pairing Type D described above.

If at the end of block j, there exists a bit at either of the transmitters that has not yet been transmitted, or any of the states $Q_{1,\mathsf{C1}}^{j-1}$, $Q_{1,\mathsf{C2}}^{j-1}$, $Q_{2,\mathsf{C1}}^{j-1}$, or $Q_{2,\mathsf{C3}}^{j-1}$ is not empty, we consider it as error type-I and halt the transmission.

Assuming that the transmission is not halted, let random variable $N_{i,\mathsf{C}_\ell}^{j}$ denote the number of bits in $Q_{i,\mathsf{C}_\ell}^{j}$, $(i,\ell) = (1,1),(1,2),(2,1),(2,3)$. Since transition of a bit to this state is distributed as independent Bernoulli RV, upon completion of block j, we have

$$\begin{aligned}
\mathbb{E}[N_{1,\mathsf{C}_1}^{j}] &= \frac{\left(q^2/p^2\right)\Pr\left(\text{Case } 1\right)}{1-\sum_{j=9,10,\ldots,16}\Pr\left(\text{Case j}\right)}m = pq^2m,\\
\mathbb{E}[N_{1,\mathsf{C}_2}^{j}] &= \frac{\Pr\left(\text{Case } 2\right)}{1-\sum_{j=9,10,\ldots,16}\Pr\left(\text{Case j}\right)}m = p^2qm,\\
\mathbb{E}[N_{2,\mathsf{C}_1}^{j}] &= \frac{\left(q^2/p^2\right)\Pr\left(\text{Case } 1\right)}{\sum_{j=2,3,4,9,10,11}\Pr\left(\text{Case j}\right)+\frac{q^2}{p^2}\sum_{j=1,12}\Pr\left(\text{Case j}\right)}2qm = pq^2m,\\
\mathbb{E}[N_{2,\mathsf{C}_3}^{j}] &= \frac{\Pr\left(\text{Case } 3\right)}{\sum_{j=2,3,4,9,10,11}\Pr\left(\text{Case j}\right)+\frac{q^2}{p^2}\sum_{j=1,12}\Pr\left(\text{Case j}\right)}2qm = p^2qm.
\end{aligned} \tag{84}$$

If the event $\left[N_{i,\mathsf{C}_\ell}^{j} \geq \mathbb{E}[N_{i,\mathsf{C}_\ell}^{j}] + m^{\frac{2}{3}}\right]$ occurs, we consider it as error type-II and we halt the transmission. At the end of block 1, we add 0's (if necessary) to $Q_{i,\mathsf{C}_\ell}^{j}$ so that the total number of bits is equal to $\mathbb{E}[N_{i,\mathsf{C}_\ell}^{j}] + m^{\frac{2}{3}}$. Using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II decreases exponentially with m.

Achievability strategy for block $b+1$: In the final communication block, transmitters do not communicate any new data bit.

If at time instant t Case 15 occurs, Tx_i sends a bit from queue $Q_{i,\mathsf{C1}}^{b}$ and removes it from the this queue. If at time instant t Case 15 occurs and $Q_{i,\mathsf{C1}}^{b}$ is empty, then Tx_i remains silent. If at time instant t Case 14 (13) or 12 occurs, Tx_1 (Tx_2) sends a bit from queue $Q_{1,\mathsf{C2}}^{b}$ ($Q_{2,\mathsf{C3}}^{b}$) and removes it from the this queue.

If at the end of block j any of the states $Q_{1,\mathsf{C1}}^{b}$, $Q_{1,\mathsf{C2}}^{b}$, $Q_{2,\mathsf{C1}}^{b}$, or $Q_{2,\mathsf{C3}}^{b}$ is not empty, we consider it as error type-I and halt the transmission.

Note that if the transmission is not halted, any bit is either available at its intended receiver interference-free, or the interfering bits is provided to the receiver in the following block. The probability that the transmission strategy halts at the end of each block can be bounded by the summation of error probabilities of types I and II. Using Chernoff-Hoeffding bound, we can show that the probability that the transmission strategy halts at any point approaches zero for $m \to \infty$.

Now, since each block has $n = m/p + \left(2/q^4\right)m^{2/3}$ time instants and the probability that the transmission strategy halts at any point approaches zero for $m \to \infty$, we achieve a rate tuple

$$\frac{b}{b+1}\left(p, 2pq\right), \tag{85}$$

as $m \to \infty$. Finally letting $b \to \infty$, we achieve the desired rate tuple.

Appendix B
Converse Proof of Theorem 1 [Instantaneous-CSIT]

The derivation of the outer-bound on individual rates is simple, however for the completeness of the results, we include the proof here. This outer-bound can be used for other theorems as needed. To derive the outer-bound on

R_1, we have

$$\begin{aligned}
nR_1 &= H(W_1) \overset{(a)}{=} H(W_1|G^n) \\
&\overset{(b)}{=} H(W_1|X_2^n, G^n) \\
&\overset{\text{(Fano)}}{\leq} I(W_1; Y_1^n|X_2^n, G^n) + n\epsilon_n \\
&\overset{\text{(data proc.)}}{\leq} I(X_1^n; Y_1^n|X_2^n, G^n) + n\epsilon_n \\
&= H(Y_1^n|X_2^n, G^n) - H(Y_1^n|X_1^n, X_2^n, G^n) + n\epsilon_n \\
&= H(G_{11}^n X_1^n|X_2^n, G^n) + n\epsilon_n \\
&\leq pn + n\epsilon_n,
\end{aligned} \tag{86}$$

where $\epsilon_n \to 0$ as $n \to \infty$; (a) holds since message W_1 is independent of G^n; and (b) holds since given G^n, W_1 is independent of X_2^n, see (44). Similarly, we have

$$nR_2 \leq pn + n\epsilon_n. \tag{87}$$

dividing both sides by n and let $n \to \infty$, we have

$$\begin{cases} R_1 \leq p \\ R_2 \leq p \end{cases} \tag{88}$$

The outer-bound on $R_1 + R_2$ follows from the proof of Theorem 4 in Section X.

Appendix C
Achievability Proof of Theorem 2: Sum-rate for $0 \leq p < 0.5$

In this appendix, we provide the achievability proof of Theorem 2 with Delayed-CSIT and for $0 \leq p < 0.5$. We provide the an achievability strategy for rate tuple

$$R_1 = R_2 = \min \left\{ p, \frac{(1-q^2)}{1 + (1-q^2)^{-1} p} \right\}. \tag{89}$$

Let the messages of transmitters one and two be denoted by $\mathrm{W}_1 = a_1, a_2, \ldots, a_m$, and $\mathrm{W}_2 = b_1, b_2, \ldots, b_m$, respectively, where a_i's and b_i's are picked uniformly and independently from $\{0, 1\}$, $i = 1, \ldots, m$. We show that it is possible to communicate these bits in

$$n = \max \left\{ m/p, \left(1-q^2\right)^{-1} m + \left(1-q^2\right)^{-2} pm \right\} + O\left(m^{2/3}\right) \tag{90}$$

time instants with vanishing error probability (as $m \to \infty$). Therefore achieving the rates given in (89) as $m \to \infty$.

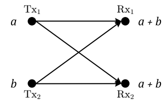
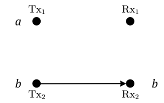
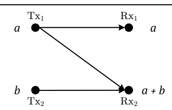
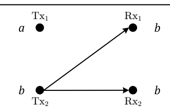
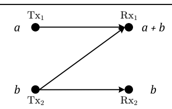
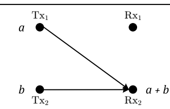
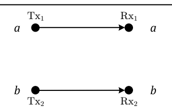
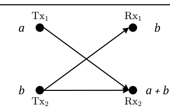
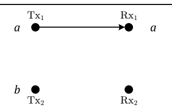
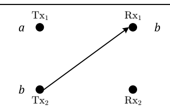
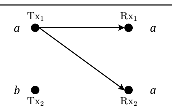
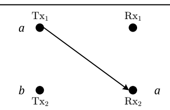
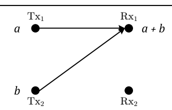
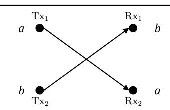
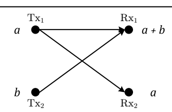
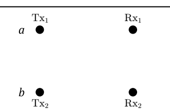
Phase 1 [uncategorized transmission]: At the beginning of the communication block, we assume that the bits at Tx_i are in queue $Q_{i \to i}$, $i = 1, 2$. At each time instant, Tx_i sends out a bit from $Q_{i \to i}$ and this bit will either stay in the initial queue or a transition to a new queue will take place. Table III summarizes the transitions for each channel realization. The arguments are very similar to our discussion in Section V, and the only difference is the way we handle Cases $7, 8, 11$, and 12. We provide some details about these cases.

For Cases 7 ($\nearrow$) and 8 ($\boxtimes$), in Section V, we updated the status of the transmitted bit of Tx_2 to $Q_{2 \to \{1,2\}}$. However, this scheme is suboptimal for $0 \leq p < 0.5$, and instead we update the status of the transmitted bit of Tx_2 to an intermediate queue $Q_{2,INT}$. Then in Phase 2, we retransmit these bits and upgrade their status once more. Similar story holds for Cases 11 and 12. The main reason for doing this is as follows. As we discussed in Section IV, there are many opportunities to combine bits in order to improve the achievable rates. However, we could never combine the bits that were transmitted in Cases $7, 8, 11$, or 12 with other bits. This was not an issue for $0.5 \leq p \leq 1$, however for $0 \leq p < 0.5$, we need to find a way to combine these bits with other bits in future time instants. To do so, the only way is to keep them in an intermediate queue and retransmit them again in Phase 2.

Phase 1 goes on for

$$\left(1-q^2\right)^{-1} m + m^{\frac{2}{3}} \tag{91}$$

TABLE III
SUMMARY OF PHASE 1 FOR THE ACHIEVABILITY SCHEME OF CORNER POINT B. BIT "a" REPRESENTS A BIT IN $Q_{1\rightarrow 1}$ WHILE BIT "b" REPRESENTS A BIT IN $Q_{2\rightarrow 2}$.

case ID	channel realization at time instant n	state transition	case ID	channel realization at time instant n	state transition
1		$\begin{cases} a \rightarrow Q_{1,\mathsf{C}_1} \\ b \rightarrow Q_{2,\mathsf{C}_1} \end{cases}$	9		$\begin{cases} a \rightarrow Q_{1\rightarrow 1} \\ b \rightarrow Q_{2,F} \end{cases}$
2		$\begin{cases} a \rightarrow Q_{1\rightarrow 2\vert 1} \\ b \rightarrow Q_{2,F} \end{cases}$	10		$\begin{cases} a \rightarrow Q_{1\rightarrow 1} \\ b \rightarrow Q_{2,F} \end{cases}$
3		$\begin{cases} a \rightarrow Q_{1,F} \\ b \rightarrow Q_{2\rightarrow 1\vert 2} \end{cases}$	11		$\begin{cases} a \rightarrow Q_{1,INT} \\ b \rightarrow Q_{2,F} \end{cases}$
4		$\begin{cases} a \rightarrow Q_{1,F} \\ b \rightarrow Q_{2,F} \end{cases}$	12		$\begin{cases} a \rightarrow Q_{1,INT} \\ b \rightarrow Q_{2,F} \end{cases}$
5		$\begin{cases} a \rightarrow Q_{1,F} \\ b \rightarrow Q_{2\rightarrow 2} \end{cases}$	13		$\begin{cases} a \rightarrow Q_{1\rightarrow 1} \\ b \rightarrow Q_{2\rightarrow 2\vert 1} \end{cases}$
6		$\begin{cases} a \rightarrow Q_{1,F} \\ b \rightarrow Q_{2\rightarrow 2} \end{cases}$	14		$\begin{cases} a \rightarrow Q_{1\rightarrow 1\vert 2} \\ b \rightarrow Q_{2\rightarrow 2} \end{cases}$
7		$\begin{cases} a \rightarrow Q_{1,F} \\ b \rightarrow Q_{2,INT} \end{cases}$	15		$\begin{cases} a \rightarrow Q_{1\rightarrow 1\vert 2} \\ b \rightarrow Q_{2\rightarrow 2\vert 1} \end{cases}$
8		$\begin{cases} a \rightarrow Q_{1,F} \\ b \rightarrow Q_{2,INT} \end{cases}$	16		$\begin{cases} a \rightarrow Q_{1\rightarrow 1} \\ b \rightarrow Q_{2\rightarrow 2} \end{cases}$

time instants and if at the end of this phase either of the queues $Q_{i\rightarrow i}$ is not empty, we declare error type-I and halt the transmission.

Assuming that the transmission is not halted, upon completion of Phase 1, we have

$$\begin{aligned}
\mathbb{E}[N_{1,\mathsf{C}_1}] &= \frac{\Pr\left(\text{Case 1}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)}m = (1-q^2)^{-1}p^4 m, \\
\mathbb{E}[N_{1\rightarrow 2|1}] &= \frac{\Pr\left(\text{Case 2}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)}m = (1-q^2)^{-1}p^3 q m, \\
\mathbb{E}[N_{1\rightarrow 1|2}] &= \frac{\sum_{j=14,15}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)}m = (1-q^2)^{-1}p q^2 m, \\
\mathbb{E}[N_{1,INT}] &= \frac{\sum_{j=11,12}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)}m = (1-q^2)^{-1}p^2 q m, \qquad (92)
\end{aligned}$$

similarly, we have

$$\begin{aligned}
\mathbb{E}[N_{2,\mathsf{C}_1}] &= \frac{\Pr\left(\text{Case 1}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)}m = (1-q^2)^{-1}p^4m, \\
\mathbb{E}[N_{2\rightarrow 1|2}] &= \frac{\Pr\left(\text{Case 3}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)}m = (1-q^2)^{-1}p^3qm, \\
\mathbb{E}[N_{2\rightarrow 2|1}] &= \frac{\sum_{j=13,15}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)}m = (1-q^2)^{-1}pq^2m, \\
\mathbb{E}[N_{2,INT}] &= \frac{\sum_{j=7,8}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)}m = (1-q^2)^{-1}p^2qm,
\end{aligned} \tag{93}$$

If the event $\left[N \geq \mathbb{E}[N] + m^{\frac{2}{3}}\right]$ for $N = N_{i,\mathsf{C}_1}, N_{i\rightarrow i|\bar{i}}, N_{i\rightarrow \bar{i}|i}, N_{i,INT}$, $i=1,2$, occurs, we consider it as error type-II and we halt the transmission strategy. At the end of Phase 1, we add 0's (if necessary) in order to make queues Q_{i,C_1}, $Q_{i\rightarrow j|\bar{j}}$, and $Q_{i,INT}$ of size equal to $\mathbb{E}[N_{i,\mathsf{C}_1}]+m^{\frac{2}{3}}$, $\mathbb{E}[N_{i\rightarrow j|\bar{j}}]+m^{\frac{2}{3}}$, and $\mathbb{E}[N_{i,INT}]+m^{\frac{2}{3}}$ respectively, $i=1,2$, and $j=i,\bar{i}$. Furthermore, using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II decreases exponentially with m.

Phase 2 [upgrading status of interfering bits in Q_{1,C_j}]: In this phase, we focus on the bits in $Q_{1,INT}$ and $Q_{2,INT}$. At each time instant, Tx_i picks a bit from $Q_{i,INT}$ and sends it. This bit will either stay in $Q_{i,INT}$ or a transition to a new queue will take place. Table IV describes what happens to the status of the bits if either of the 16 possible cases occurs. Due to symmetry, we only describe the transitions for bits in $Q_{1,INT}$. Consider bit "a" in $Q_{1,INT}$.

- Cases $1,2,3,4$, and 5: The transitions for these cases are consistent with our previous discussions.
- Cases $9,10,11,12,13$, and 16: In these cases, it is easy to see that no change occurs in the status of bit a.
- Case 6: In this case, bit a is delivered to both receivers and hence, no further transmission is required. Therefore, it joins $Q_{1,F}$.
- Case 7: Here with slight abuse of, Q_{1,C_1} represents the bits of Tx_1 that are received at both receivers with interference but not necessarily in Case 1, $i=1,2$. For instance if at a given time, Tx_1 sends a bit from $Q_{1,INT}$ and Case 7 occurs, then this bit joins Q_{1,C_1} since now both receivers have received this bit with interference.
- Case 8: In this case, bit a is available at Rx_2 but it is interfered at Rx_1 by bit b. However, in Case 8 no change occurs for the bits in $Q_{2,INT}$. Therefore, since bit b will be retransmitted until it is provided to Rx_1, no retransmission is required for bit a and it joins $Q_{1,F}$.
- Cases 14 and 15: If either of these cases occur, bit a becomes available at Rx_2 and is needed at Rx_1. Thus, we update the status of such bits to $Q_{1\rightarrow 1|2}$.

Phase 2 goes on for

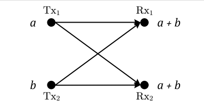
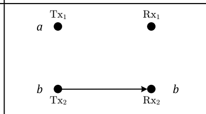
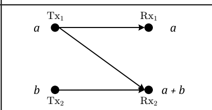
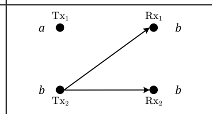
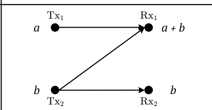
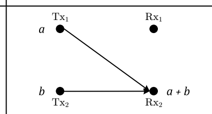
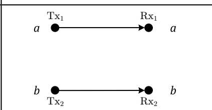
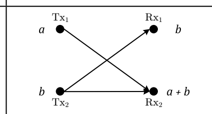
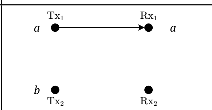
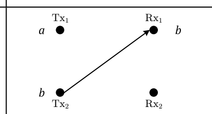
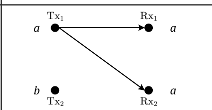
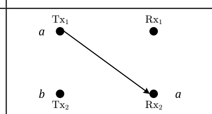
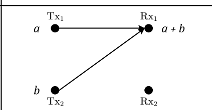
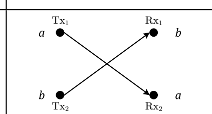
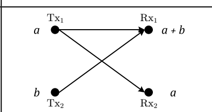
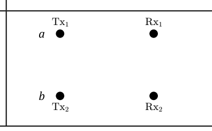
$$\begin{aligned}
&\left(1-\sum_{i=9,10,11,12,13,16}\Pr\left(\text{Case } i\right)\right)^{-1}\left(1-q^2\right)^{-1}p^2qm + 2m^{\frac{2}{3}} \\
&\quad = \left(1-\left[p^2q+q^2\right]\right)^{-1}\left(1-q^2\right)^{-1}p^2qm + 2m^{\frac{2}{3}}
\end{aligned} \tag{94}$$

time instants and if at the end of this phase either of the states $Q_{i,INT}$ is not empty, we declare error type-I and halt the transmission.

Assuming that the transmission is not halted, upon completion of Phase 2, the states $Q_{1,INT}$ and $Q_{2,INT}$ are empty and we have

$$\begin{aligned}
\mathbb{E}[N_{i,\mathsf{C}_1}] &= (1-q^2)^{-1}p^4m + \frac{\sum_{j=1,7}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,11,12,13,16}\Pr\left(\text{Case i}\right)}(p^2qm+m^{2/3}) \\
&= (1-q^2)^{-1}p^4m + \left(1-\left[p^2q+q^2\right]\right)^{-1}\left(p^4+p^2q^2\right)\left(p^2qm+m^{2/3}\right),
\end{aligned} \tag{95}$$

TABLE IV
SUMMARY OF PHASE 2 FOR THE ACHIEVABILITY SCHEME OF CORNER POINT B. BIT "a" REPRESENTS A BIT IN $Q_{1,INT}$ WHILE BIT "b" REPRESENTS A BIT IN $Q_{2,INT}$.

case ID	channel realization at time instant n	state transition	case ID	channel realization at time instant n	state transition
1		$\begin{cases} a \to Q_{1,\mathsf{C}_1} \\ b \to Q_{2,\mathsf{C}_1} \end{cases}$	9		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2\to1\|2} \end{cases}$
2		$\begin{cases} a \to Q_{1\to2\|1} \\ b \to Q_{2,F} \end{cases}$	10		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2,F} \end{cases}$
3		$\begin{cases} a \to Q_{1,F} \\ b \to Q_{2\to1\|2} \end{cases}$	11		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2,\mathsf{C}_1} \end{cases}$
4		$\begin{cases} a \to Q_{1\to2\|1} \\ b \to Q_{2\to1\|2} \end{cases}$	12		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2,F} \end{cases}$
5		$\begin{cases} a \to Q_{1\to2\|1} \\ b \to Q_{2,INT} \end{cases}$	13		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2\to2\|1} \end{cases}$
6		$\begin{cases} a \to Q_{1,F} \\ b \to Q_{2,INT} \end{cases}$	14		$\begin{cases} a \to Q_{1\to1\|2} \\ b \to Q_{2,INT} \end{cases}$
7		$\begin{cases} a \to Q_{1,\mathsf{C}_1} \\ b \to Q_{2,INT} \end{cases}$	15		$\begin{cases} a \to Q_{1\to1\|2} \\ b \to Q_{2\to2\|1} \end{cases}$
8		$\begin{cases} a \to Q_{1,F} \\ b \to Q_{2,INT} \end{cases}$	16		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2,INT} \end{cases}$

similarly, we have

$$\begin{aligned}\mathbb{E}[N_{i\to\bar{i}|i}] &= (1-q^2)^{-1}p^3qm + \frac{\sum_{j=2,4,5}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,11,12,13,16}\Pr\left(\text{Case i}\right)}(p^2qm+m^{2/3}) \\ &= (1-q^2)^{-1}p^3qm + \left(1-\left[p^2q+q^2\right]\right)^{-1}\left(p^2q+p^2q^2\right)\left(p^2qm+m^{2/3}\right), \end{aligned} \tag{96}$$

and

$$\begin{aligned}\mathbb{E}[N_{i\to i|\bar{i}}] &= (1-q^2)^{-1}pq^2m + \frac{\sum_{j=14,15}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,11,12,13,16}\Pr\left(\text{Case i}\right)}(p^2qm+m^{2/3}) \\ &= (1-q^2)^{-1}pq^2m + \left(1-\left[p^2q+q^2\right]\right)^{-1}pq^2\left(p^2qm+m^{2/3}\right). \end{aligned} \tag{97}$$

If the event $\left[N \geq \mathbb{E}[N] + m^{\frac{2}{3}}\right]$ for $N = N_{i,\mathsf{C}_1}, N_{i\to i|\bar{i}}, N_{i\to\bar{i}|i}$, $i=1,2$, occurs, we consider it as error type-II and we halt the transmission strategy. At the end of Phase 2, we add 0's (if necessary) in order to make queues Q_{i,C_1} and $Q_{i\to j|\bar{j}}$ of size equal to $\mathbb{E}[N_{i,\mathsf{C}_1}] + m^{\frac{2}{3}}$, and $\mathbb{E}[N_{i\to j|\bar{j}}] + m^{\frac{2}{3}}$ respectively, $i=1,2$, and $j=i,\bar{i}$. Using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II decreases exponentially with m.

Phase 3 [encoding and retransmission]: In this phase, Tx_i communicates the bits in $Q_{i\to i|\bar{i}}$ to Rx_i, $i=1,2$. However, it is possible to create XOR of these bits with the bits in $Q_{i\to\bar{i}|i}$ and the bits in Q_{i,C_1} to create bits of common

interest. To do so, we first encode the bits in these states using the results of [27], and then we create the XOR of the encoded bits.

In other words, given $\epsilon, \delta > 0$, Tx_i encodes all the bits in $Q_{i\to i|\bar{i}}$ at rate $p-\delta$ using random coding scheme of [27]. Similarly, Tx_i encodes $q\left(\mathbb{E}[N_{i\to i|\bar{i}}]+m^{\frac{2}{3}}\right)$ bits from $Q_{i\to \bar{i}|i}$ and Q_{i,C_1} at rate $pq-\delta$ (if there are less bits in $Q_{i\to \bar{i}|i}$ and Q_{i,C_1}, then encode all of the bits in these queues). More precisely, first Tx_i encodes bits from $Q_{i\to \bar{i}|i}$, and if the number of bits in $Q_{i\to \bar{i}|i}$ is less than $q\left(\mathbb{E}[N_{i\to i|\bar{i}}]+m^{\frac{2}{3}}\right)$, then Tx_i uses bits in Q_{i,C_1}. Tx_i will then communicate the XOR of these encoded bits.

Note that since Rx_i already knows the bits in $Q_{i\to \bar{i}|i}$, it can remove the corresponding part of the received signal. Then since the channel from $\mathsf{Tx}_{\bar{i}}$ to Rx_i can be viewed as a binary erasure channel with success probability of pq, from [27], we know that Rx_i can decode $Q_{\bar{i},\mathsf{C}_1}$ with decoding error probability less than or equal to ϵ. Thus, Rx_i can decode the transmitted bits from $Q_{\bar{i},\mathsf{C}_1}$ and use them to decode the bits in Q_{i,C_1}. Then, Rx_i removes the contribution of the bits in Q_{i,C_1} the received signal. Finally, since the channel from Tx_i to Rx_i can be viewed as a binary erasure channel with success probability of p, from [27], we know that Rx_i can decode $Q_{i\to i|\bar{i}}$ with decoding error probability less than or equal to ϵ.

If an error occurs in decoding of the encoded bits, we halt the transmission. Assuming that the transmission is not halted, at the end of Phase 3, $Q_{i\to i|\bar{i}}$ becomes empty and there are

$$\left(\mathbb{E}[N_{i\to \bar{i}|i}]+\mathbb{E}[N_{i,\mathsf{C}_1}]+2m^{\frac{2}{3}}-q\left(\mathbb{E}[N_{i\to i|\bar{i}}]+m^{\frac{2}{3}}\right)\right)^{+} \tag{98}$$

bits left in $Q_{i\to \bar{i}|i}$ and Q_{i,C_1}.

If $Q_{i\to \bar{i}|i}$ and Q_{i,C_1} are also empty, the transmission strategy ends here. Otherwise, we merge the remaining bits in $Q_{i\to \bar{i}|i}$ (if any) with the bits in Q_{i,C_1} as **Type-III** (see Section IV) and put the XOR of them in $Q_{i\to\{1,2\}}$, $i=1,2$. Finally, we need to describe what happens to the remaining bits in Q_{i,C_1}. As mentioned before, a bit in Q_{i,C_1} can be viewed as a bit of common interest by itself. For the remaining bits in Q_{1,C_1}, we put the first half in $Q_{1\to\{1,2\}}$ (suppose m is picked such that the remaining number of bits is even). Note that if these bits are delivered to Rx_2, then Rx_2 can decode the first half of the bits in Q_{2,C_1}. Therefore, the first half of the bits in Q_{2,C_1} join $Q_{2,F}$.

Phase 4 [communicating bits of common interest]: During Phase 4, we deliver the bits in $Q_{1\to\{1,2\}}$ and $Q_{2\to\{1,2\}}$ using the transmission strategy for the two-source multicast problem. More precisely, the bits in $Q_{i\to\{1,2\}}$ will be considered as the message of transmitter Tx_i and they will be encoded as in the achievability scheme of Lemma 2, $i=1,2$. Fix $\epsilon,\delta>0$, from Lemma 2 we know that the rate tuple

$$(R_1,R_2)=\frac{1}{2}\left((1-q^2)-\delta,(1-q^2)-\delta\right)$$

is achievable with decoding error probability less than or equal to ϵ. Thus, using Lemma 2, we can communicate the remaining bits at rate $(1-q^2)-\delta$ with decoding error probability less than or equal to ϵ. If an error occurs in decoding of the encoded bits, we halt the transmission.

Using Chernoff-Hoeffding bound and the results of [27], we can show that the probability that the transmission strategy halts at any point approaches zero for $\epsilon,\delta\to 0$ and $m\to\infty$. Moreover, it is easy to verify that for $0\le p\le\left(3-\sqrt{5}\right)/2$, at the end of Phase 3, $Q_{i\to \bar{i}|i}$ and Q_{i,C_1} are empty and the transmission strategy ends there. However, for $\left(3-\sqrt{5}\right)/2<p<0.5$, the transmission strategy continues to Phase 4. Therefore, we can show that if no error occurs, the transmission strategy end in

$$n=\max\left\{m/p,\left(1-q^2\right)^{-1}m+\left(1-q^2\right)^{-2}pm\right\}+O\left(m^{2/3}\right) \tag{99}$$

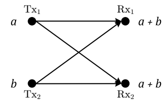
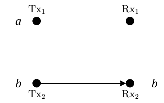
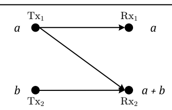
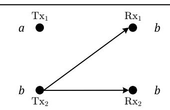
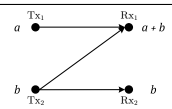
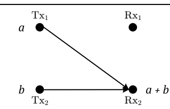
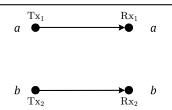
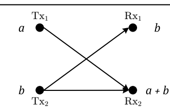
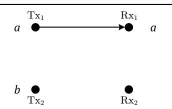
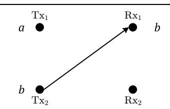
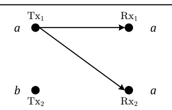
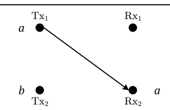
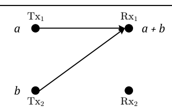
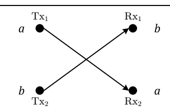
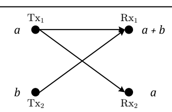
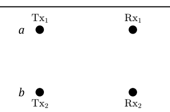
time instants. Therefore achieving the rates given in (89).

Appendix D
Achievability Proof of Theorem 2: Corner Point C

In this appendix, we describe the achievability strategy for corner point C depicted in Figure 17(b), *i.e.*

$$(R_1,R_2)=(pq(1+q),p)\,. \tag{100}$$

TABLE V
SUMMARY OF PHASE 1 FOR THE ACHIEVABILITY SCHEME OF CORNER POINT C. BIT “a” REPRESENTS A BIT IN $Q_{1\to 1}$ WHILE BIT “b” REPRESENTS A BIT IN $Q_{2\to 2}$.

case ID	channel realization at time instant n	state transition	case ID	channel realization at time instant n	state transition
1		$\begin{cases} a \to Q_{1,\mathsf{C}_1} \\ b \to Q_{2,F} \end{cases}$	9		$\begin{cases} a \to Q_{1\to 1} \\ b \to Q_{2,F} \end{cases}$
2		$\begin{cases} a \to Q_{1,OP} \\ b \to Q_{2,F} \end{cases}$	10		$\begin{cases} a \to Q_{1\to 1} \\ b \to Q_{2,F} \end{cases}$
3		$\begin{cases} a \to Q_{1,F} \\ b \to Q_{2,OP} \end{cases}$	11		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2,F} \end{cases}$
4		$\begin{cases} a \to Q_{1,F} \\ b \to Q_{2,F} \end{cases}$	12		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2,F} \end{cases}$
5		$\begin{cases} a \to Q_{1,F} \\ b \to Q_{2\to 2} \end{cases}$	13		$\begin{cases} a \to Q_{1\to 1} \\ b \to Q_{2\to 2\vert 1} \end{cases}$
6		$\begin{cases} a \to Q_{1,F} \\ b \to Q_{2\to 2} \end{cases}$	14		$\begin{cases} a \to Q_{1\to 1\vert 2} \\ b \to Q_{2\to 2} \end{cases}$
7		$\begin{cases} a \to Q_{1,F} \\ b \to Q_{2,INT} \end{cases}$	15		$\begin{cases} a \to Q_{1\to 1\vert 2} \\ b \to Q_{2\to 2\vert 1} \end{cases}$
8		$\begin{cases} a \to Q_{1,F} \\ b \to Q_{2,INT} \end{cases}$	16		$\begin{cases} a \to Q_{1\to 1} \\ b \to Q_{2\to 2} \end{cases}$

Let the messages of transmitters one and two be denoted by $\mathrm{W}_1 = a_1, a_2, \ldots, a_{m_1}$, and $\mathrm{W}_2 = b_1, b_2, \ldots, b_m$, respectively, where data bits a_i's and b_i's are picked uniformly and independently from $\{0, 1\}$, and $m_1 = q(1+q)m$ (suppose the parameters are such that $m, m_1 \in \mathbb{Z}$). Note that for $\left(3 - \sqrt{5}\right)/2 < p \leq 1$, we have $q(1+q) < 1$. We show that it is possible to communicate these bits in

$$n = \frac{1}{p}m + O\left(m^{2/3}\right) \tag{101}$$

time instants with vanishing error probability (as $m \to \infty$). Therefore, achieving corner point C as $m \to \infty$. Our transmission strategy consists of five phases as described before.

Phase 1 [uncategorized transmission]: This phase is similar to Phase 1 of the achievability strategy for the optimal sum-rate point A. The main difference is due to the fact that the transmitters start with unequal number of bits. At the beginning of the communication block, we assume that the bits $a_1, a_2, \ldots, a_{m_1}$ at Tx_1 and the bits $b_1, b_2, \ldots, b_{m_1}$ at Tx_2 are in queues $Q_{1\to 1}$ and $Q_{2\to 2}$ respectively.

Remark 17: Note that Tx_2 has m initial bits, however, only m_1 of them are in $Q_{2\to 2}$ at the beginning of the communication block.

At each time instant t, Tx_i sends out a bit from $Q_{i\to i}$, and this bit will either stay in the initial queues or a transition will take place. Based on the channel realizations, a total of 16 possible configurations may occur at any time instant. Table V summarizes the transition from the initial queue for each channel realization.

In comparison to the achievability strategy of the sum-rate point A, we have new queues for the bits:

1) $Q_{i,OP}$ denotes the bits that have caused interference at the unintended receiver and this interference has to get resolved.
2) $Q_{1,INT}$ denotes an intermediate queue of the bits at Tx_1 that were transmitted when channel realizations 11 or 12 occurred.
3) $Q_{i,INT}$ denotes an intermediate queue of the bits at Tx_2 that were transmitted when channel realizations 7 or 8 occurred.

Phase 1 goes on for

$$(1/p-1)\,m+m^{\frac{2}{3}} \tag{102}$$

time instants and if at the end of this phase, either of queues $Q_{i\rightarrow i}$ is not empty, we declare error type-I and halt the transmission.

Assuming that the transmission is not halted, let random variable N_{1,C_1}, $N_{i,OP}$, $N_{i\rightarrow i|\bar{i}}$, and $N_{i,INT}$ denote the number of bits in Q_{1,C_1}, $Q_{i,OP}$, $Q_{i\rightarrow i|\bar{i}}$, and $Q_{i,INT}$ respectively $i=1,2$. The transmission strategy will be halted and an error (that we refer to as error type-II) will occur, if any of the following events happens.

$$N_{1,\mathsf{C}_1} > \mathbb{E}[N_{1,\mathsf{C}_1}]+m^{\frac{2}{3}} \triangleq n_{1,\mathsf{C}_1}; \quad N_{i,OP} > \mathbb{E}[N_{i,OP}]+m^{\frac{2}{3}} \triangleq n_{i,OP}, \quad i=1,2; N_{i\rightarrow i|\bar{i}} > \mathbb{E}[N_{i\rightarrow i|\bar{i}}]+m^{\frac{2}{3}} \triangleq n_{i\rightarrow i|\bar{i}}, \quad i=1, \tag{103}$$

From basic probability, and we have

$$\begin{aligned}
\mathbb{E}[N_{1,\mathsf{C}_1}] &= \frac{\Pr\left(\text{Case 1}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m_1 = (1-q^2)^{-1}p^4 m_1 = p^3 q m, \\
\mathbb{E}[N_{1,OP}] &= \frac{\Pr\left(\text{Case 2}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m_1 = (1-q^2)^{-1}p^3 q m_1 = p^2 q^2 m, \\
\mathbb{E}[N_{2,OP}] &= \frac{\Pr\left(\text{Case 3}\right)}{1-\sum_{i=5,6,14,16}\Pr\left(\text{Case i}\right)} m_1 = (1-q^2)^{-1}p^3 q m_1 = p^2 q^2 m, \\
\mathbb{E}[N_{i\rightarrow i|\bar{i}}] &= \frac{\sum_{j=14,15}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m_1 = (1-q^2)^{-1}\left(pq^3+p^2q^2\right) m_1 = q^3 m, \qquad i=1,2, \\
\mathbb{E}[N_{i,INT}] &= \frac{\sum_{j=11,12}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,13,16}\Pr\left(\text{Case i}\right)} m_1 = (1-q^2)^{-1}\left(p^3q+p^2q^2\right) m_1 = pq^2 m, \qquad i=1,2.
\end{aligned} \tag{104}$$

Furthermore, using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II decreases exponentially with m.

At the end of Phase 1, we add 0's (if necessary) in order to make queues Q_{1,C_1}, $Q_{i,OP}$, $Q_{i\rightarrow i|\bar{i}}$, and $Q_{i,INT}$ of size equal to n_{1,C_1}, $n_{i,OP}$, $n_{i\rightarrow i|\bar{i}}$, and $n_{i,INT}$ respectively as defined in (103), $i=1,2$. For the rest of this appendix, we assume that Phase 1 is completed and no error has occurred.

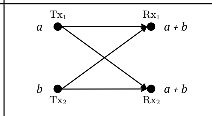
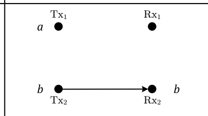
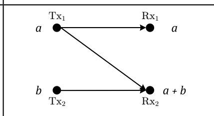
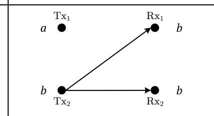
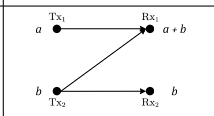
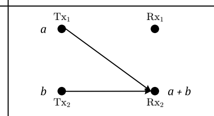
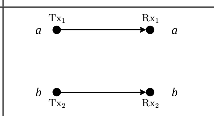
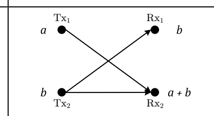
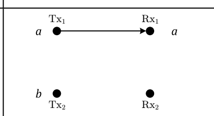
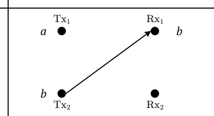
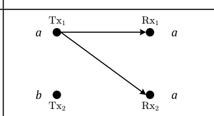
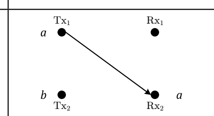
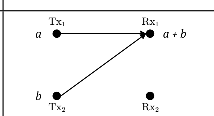
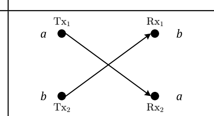
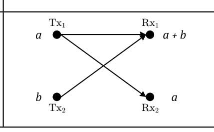
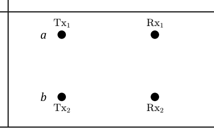
Phase 2 [updating status of the bits in $Q_{i,INT}$]: In this phase, we focus on the bits in $Q_{i,INT}$, $i=1,2$. The ultimate goal is to deliver the bits in $Q_{i,INT}$ to both receivers. At each time instant, Tx_i picks a bit from $Q_{i,INT}$ and sends it. This bit will either stay in $Q_{i,INT}$ or a transition to a new queue will take place. Table VI describes what happens to the status of the bits if either of the 16 cases occurs.

Here, we describe what happens to the status of a bit in $Q_{1,INT}$ if either of the 16 channel realizations occur. The description for a bit in $Q_{2,INT}$ is very similar and is summarized in Table VI. Consider a bit "a" in $Q_{1,INT}$. At each time instant, 16 possible cases may occur:

- Cases 9,10,11,12,13, and 16: In these cases, it is easy to see that no change occurs in the status of bit a.
- Case 6: In this case, bit a is delivered to both receivers and hence, no further transmission is required. Therefore, it joins $Q_{1,F}$.
- Case 8: In this case, bit a is available at Rx_2 but it is interfered at Rx_1 by bit b. However, in Case 8 no change occurs for the bits in $Q_{2,INT}$. Therefore, since bit b will be retransmitted until it is provided to Rx_1, no retransmission is required for bit a and it joins $Q_{1,F}$.
- Cases 14 and 15: If either of these cases occur, bit a becomes available at Rx_2 and is needed at Rx_1. Thus, we update the status of such bits to $Q_{1\rightarrow 1|2}$.

TABLE VI

SUMMARY OF PHASE 2 FOR THE ACHIEVABILITY SCHEME OF CORNER POINT B. BIT "a" REPRESENTS A BIT IN $Q_{1,INT}$ WHILE BIT "b" REPRESENTS A BIT IN $Q_{2,INT}$.

case ID	channel realization at time instant n	state transition	case ID	channel realization at time instant n	state transition
1		$\begin{cases} a \to Q_{1,\mathsf{C}_1} \\ b \to Q_{2,F} \end{cases}$	9		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2,OP} \end{cases}$
2		$\begin{cases} a \to Q_{1,OP} \\ b \to Q_{2,OP} \end{cases}$	10		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2,F} \end{cases}$
3		$\begin{cases} a \to Q_{1,\mathsf{C}_1} \\ b \to Q_{2,OP} \end{cases}$	11		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2,OP} \end{cases}$
4		$\begin{cases} a \to Q_{1,OP} \\ b \to Q_{2,OP} \end{cases}$	12		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2,F} \end{cases}$
5		$\begin{cases} a \to Q_{1,OP} \\ b \to Q_{2,INT} \end{cases}$	13		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2\to 2\|1} \end{cases}$
6		$\begin{cases} a \to Q_{1,F} \\ b \to Q_{2,INT} \end{cases}$	14		$\begin{cases} a \to Q_{1\to 1\|2} \\ b \to Q_{2,INT} \end{cases}$
7		$\begin{cases} a \to Q_{1,OP} \\ b \to Q_{2,INT} \end{cases}$	15		$\begin{cases} a \to Q_{1\to 1\|2} \\ b \to Q_{2\to 2\|1} \end{cases}$
8		$\begin{cases} a \to Q_{1,F} \\ b \to Q_{2,INT} \end{cases}$	16		$\begin{cases} a \to Q_{1,INT} \\ b \to Q_{2,INT} \end{cases}$

- Cases 1,2,3,4,5, and 7: If either of these cases occur, we upgrade the status of bit a to the opportunistic state $Q_{1,OP}$, meaning that from now on bit a has to be provided to either Rx_2 or both receivers such that it causes no further interference. For instance, if Case 2 occurs, providing bit a to both receivers is suffiecient to decode the simultaneously transmitted bits.

Phase 2 goes on for

$$\left(1 - \left[p^3 q + 2pq^2 + q^4\right]\right)^{-1} pq^2 m + 2m^{\frac{2}{3}} \tag{105}$$

time instants, and if at the end of this phase either of the states $Q_{i,INT}$ is not empty, we declare error type-I and halt the transmission.

Assuming that the transmission is not halted, since transition of a bit to this state is distributed as independent

Bernoulli RV, upon completion of Phase 2, we have

$$\begin{aligned}
\mathbb{E}[N_{1,\mathsf{C}_1}] &= p^3qm + \frac{\sum_{j=1,3}\Pr\left(\text{Case j}\right)}{1-\sum_{i=9,10,11,12,13,16}\Pr\left(\text{Case i}\right)}(pq^2m+m^{2/3}) \\
&= p^3qm + \left(1-\left[p^3q+2pq^2+q^4\right]\right)^{-1}p^3(pq^2m+m^{2/3}), \\
\mathbb{E}[N_{1,OP}] &= p^2q^2m + \left(1-\left[p^3q+2pq^2+q^4\right]\right)^{-1}pq(pq^2m+m^{2/3}), \\
\mathbb{E}[N_{2,OP}] &= p^2q^2m + \left(1-\left[p^3q+2pq^2+q^4\right]\right)^{-1}pq(1+p^2)(pq^2m+m^{2/3}), \\
\mathbb{E}[N_{i\rightarrow i|\bar{i}}] &= q^3m + \left(1-\left[p^3q+2pq^2+q^4\right]\right)^{-1}pq^2\left(pq^2+m^{2/3}\right), \qquad i=1,2.
\end{aligned} \tag{106}$$

The transmission strategy will halt and an error (that we refer to as error type-II) will occur, if any of the following events happens.

$$N_{1,\mathsf{C}_1} > \mathbb{E}[N_{1,\mathsf{C}_1}] + m^{\frac{2}{3}} \stackrel{\triangle}{=} n_{1,\mathsf{C}_1}; \quad N_{i,OP} > \mathbb{E}[N_{i,OP}] + m^{\frac{2}{3}} \stackrel{\triangle}{=} n_{i,OP}, \quad i=1,2; N_{i\rightarrow i|\bar{i}} > \mathbb{E}[N_{i\rightarrow i|\bar{i}}] + m^{\frac{2}{3}} \stackrel{\triangle}{=} n_{i\rightarrow i|\bar{i}}, \quad i=1, \tag{107}$$

Using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II decreases exponentially with m.

Again, at the end of Phase 2, we add 0's (if necessary) in order to make queues Q_{1,C_1}, $Q_{i,OP}$, and $Q_{i\rightarrow i|\bar{i}}$ of size equal to n_{1,C_1}, $n_{i,OP}$, and $n_{i\rightarrow i|\bar{i}}$ respectively as defined in (107), $i=1,2$. For the rest of this appendix, we assume that Phase 2 is completed and no error has occurred.

Note that Tx_2 initially had m fresh data bits but during Phase 1 it only communicated m_1 of them. The rest of those bits will be transmitted during Phase 3 as described below.

Phase 3 [uncategorized transmission vs interference management]: During Phase 3, Tx_1 (the secondary user) communicates $\frac{q}{1+q}(p-q^2)m$ bits from states Q_{1,C_1} and $Q_{1,OP}$ at a rate such that both receivers can decode them at the end of Phase 3, regardless of the transmitted signal of Tx_2. In fact, at Rx_1, we have

$$\Pr\left[G_{11}[t]=1, G_{21}[t]=0\right] = pq, \tag{108}$$

and at Rx_2, we have

$$\Pr\left[G_{22}[t]=0, G_{12}[t]=1\right] = pq. \tag{109}$$

Hence, using the results of [27], we know that given any $\epsilon, \delta > 0$, Tx_1 can use a random code of rate $pq-\delta$ to encode $\frac{q}{1+q}(p-q^2)m$ bits from states Q_{1,C_1} and $Q_{1,OP}$, and transmits them such that both receivers can decode the transmitted message with error probability less than or equal to ϵ for sufficiently large block length (Tx_1 picks bits from Q_{1,C_1} and if this state becomes empty it starts picking from the bits in $Q_{1,OP}$). Since Rx_2 can decode the transmitted signal of Tx_1 in this phase, we can assume that the encoded bits of Tx_1, do not create any new interference during Phase 3.

We now describe what Tx_2 does during Phase 3. At the beginning of Phase 3, we assume that the bits $b_{m_1+1}, b_{m_1+2}, \ldots, b_m$ at Tx_2 are in state $Q_{2\rightarrow 2}$. At each time instant, Tx_2 picks a bit from $Q_{2\rightarrow 2}$ and sends it. This bit will either stay in $Q_{2\rightarrow 2}$ or a transition occurs as described below.

- Cases $1,2,3,4,9,10,11$, and 12: In these cases the direct link from Tx_2 to Rx_2 is on. Therefore, since at the end of block (assuming large enough block length), we can decode and remove the transmitted signal of Tx_1, the transmitted bit of Tx_2 leaves $Q_{2\rightarrow 2}$ and joins $Q_{2,F}$.
- Cases $7,8,13$, and 15: In these cases (assuming the transmitted signal of Tx_1 can be removed), the transmitted bit of Tx_2 becomes available at Rx_1 while it is required at Rx_2. Thus, the transmitted bit of Tx_2 leaves $Q_{2\rightarrow 2}$ and joins $Q_{2\rightarrow 2|1}$.
- Cases $5,6,14$, and 16: In these cases, no change hanppens in the status of the transmitted bit from Tx_2.

Phase 3 goes on for

$$\frac{(p-q^2)}{(1-q^2)}m + m^{\frac{2}{3}} \tag{110}$$

time instants and if at the end of this phase there is a bit left in $Q_{2\to 2}$ or an error occurs in decoding the transmitted signal of Tx_1, we declare error type-I and halt the transmission. Note that during Phase 3, the number of bits in $Q_{1\to 1|2}$ and $Q_{2,OP}$ remain unchanged.

Assuming that the transmission is not halted, since transition of a bit to this state is distributed as independent Bernoulli RV, upon completion of Phase 2, we have

$$\begin{aligned}
\mathbb{E}[N_{1,\mathsf{C}_1}] &= \left[p^3qm + \left(1 - \left[p^3q + 2pq^2 + q^4\right]\right)^{-1} p^3(pq^2m + m^{2/3}) - \frac{q}{1+q}(p-q^2)m\right]^+, \\
\mathbb{E}[N_{1,OP}] &= p^2q^2m + \left(1 - \left[p^3q + 2pq^2 + q^4\right]\right)^{-1} pq(pq^2m + m^{2/3}) \\
&\quad - \left[\frac{q}{1+q}(p-q^2)m - p^3qm - \left(1 - \left[p^3q + 2pq^2 + q^4\right]\right)^{-1} p^3(pq^2m + m^{2/3})\right]^+ .
\end{aligned} \tag{111}$$

For $\left(3-\sqrt{5}\right)/2 \leq p \leq 1$, $\mathbb{E}[N_{1,OP}]$ is non-negative. The transmission strategy will halt and an error (that we refer to as error type-II) will occur, if any of the following events happens.

1) $N_{1,\mathsf{C}_1} > \mathbb{E}[N_{1,\mathsf{C}_1}] + m^{\frac{2}{3}}$;
2) $N_{1,OP} > \mathbb{E}[N_{1,OP}] + m^{\frac{2}{3}}$;
3) $N_{2\to 2|1} > \mathbb{E}[N_{2\to 2|1}] + m^{\frac{2}{3}}$.

Using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II decreases exponentially with m.

Phase 4 [delivering interference-free bits and interference management]: In Phase 4, Tx_1 will communicate all the bits in $Q_{1\to 1|2}$. However, it is possible to create XOR of these bits with bits in $Q_{1,OP}$ in order to create bits of common interest. To do so, we first encode the bits in these states using the results of [27], and then we create the XOR of the encoded bits. On the other hand, Tx_2 will do the same to part of the bits in $Q_{2\to 2|1}$ and $Q_{2,OP}$.

More precisely, for any $\epsilon, \delta > 0$, Tx_1 encodes all the bits in $Q_{1\to 1|2}$ at rate $p-\delta$ using random coding scheme of [27]. Similarly, Tx_1 encodes

$$q^4 + \left(1 - \left[p^3q + 2pq^2 + q^4\right]\right)^{-1} p^2q^5 \tag{112}$$

bits from $Q_{1,OP}$ at rate $pq-\delta$. Then Tx_1 will communicate the XOR of these encoded bits.

During Phase 3, Tx_2 encodes same number of the bits as in $Q_{1\to 1|2}$ from $Q_{2\to 2|1}$ at rate $p-\delta$ and

$$q^4 + \left(1 - \left[p^3q + 2pq^2 + q^4\right]\right)^{-1} p^2q^5$$

bits from $Q_{2,OP}$ at rate $pq-\delta$ using random coding scheme of [27]. Then Tx_2 will communicate the XOR of these encoded bits.

Since Rx_2 already has access to the bits in $Q_{1\to 1|2}$ and $Q_{2,OP}$, it can remove their contribution from the received signals. Then for sufficiently large block length, Rx_2 can decode the transmitted bits from $Q_{1,OP}$ with decoding error probability less than or equal to ϵ. After decoding and removing this part, Rx_2 can decode the encoded bits from $Q_{2\to 2|1}$ with decoding error probability less than or equal to ϵ.

On the other hand, since Rx_1 already has access to the bits in $Q_{2\to 2|1}$ and $Q_{1,OP}$, it can remove their contribution from the received signals. Then for sufficiently large block length, Rx_1 can decode the transmitted bits from $Q_{2,OP}$ with decoding error probability less than or equal to ϵ. Finally, after decoding and removing this part, Rx_1 can decode the encoded bits from $Q_{1\to 1|2}$ with decoding error probability less than or equal to ϵ.

Phase 4 goes on for

$$\left[q^3m + \left(1 - \left[p^3q + 2pq^2 + q^4\right]\right)^{-1} pq^2\left(pq^2 + m^{2/3}\right)\right] / \left(p - \delta\right) \tag{113}$$

time instants. If an error occurs in decoding any of the encoded signals in Phase 4, we consider it as error and we halt the transmission strategy.

At the end of Phase 4, $Q_{1\to 1|2}$ becomes empty. Define

$$\beta = \left(pq - \delta\right)\left[q^3m + \left(1 - \left[p^3q + 2pq^2 + q^4\right]\right)^{-1} pq^2\left(pq^2 + m^{2/3}\right)\right] / \left(p - \delta\right), \tag{114}$$

note that as $\epsilon, \delta \to 0$, β agrees with the expression given in (112).

Upon completion of Phase 4, we have

$$\begin{aligned}\mathbb{E}[N_{1,OP}] &= p^2q^2m + \left(1 - \left[p^3q + 2pq^2 + q^4\right]\right)^{-1} pq(pq^2m + m^{2/3}) \\ &\quad - \left[\frac{q}{1+q}(p-q^2)m - p^3qm - \left(1 - \left[p^3q + 2pq^2 + q^4\right]\right)^{-1} p^3(pq^2m + m^{2/3})\right]^+ - \frac{q}{1+q}(p-q^2)m - \beta, \\ \mathbb{E}[N_{2,OP}] &= p^2q^2m + \left(1 - \left[p^3q + 2pq^2 + q^4\right]\right)^{-1} pq(1+p^2)(pq^2m + m^{2/3}) - \beta, \\ \mathbb{E}[N_{2\rightarrow 2|1}] &= \frac{pq}{(1-q^2)}\left(p - q^2\right)m. \end{aligned} \tag{115}$$

The transmission strategy will halt and an error of type-II will occur, if any of the following events happens.

1) $N_{i,OP} > \mathbb{E}[N_{i,OP}] + m^{\frac{2}{3}}$, $i = 1, 2$;
2) $N_{2\rightarrow 2|1} > \mathbb{E}[N_{2\rightarrow 2|1}] + m^{\frac{2}{3}}$.

Using Chernoff-Hoeffding bound, we can show that the probability of errors of type II decreases exponentially with m. Furthermore, the probability that an error occurs in decoding any of the encoded signals in Phase 4 can be made arbitrary small as $m \rightarrow \infty$.

Phase 5 [delivering interference-free bits and interference management]: In Phase 5, the transmitters will communicate the remaining bits in $Q_{1,OP}$, Q_{1,C_1}, $Q_{2\rightarrow 2|1}$, and $Q_{2,OP}$. Tx_1 will communicate all the bits in $Q_{1,OP}$ and Q_{1,C_1} such that for sufficiently large block length, both receivers can decode them with arbitrary small error. On the other hand, Tx_2 will communicate the bits in $Q_{2\rightarrow 2|1}$ and $Q_{2,OP}$ similar to Phase 4 with one main difference. Since both receivers can completely remove the contribution of Tx_1 at the end of the block, Tx_2 can send the bits in $Q_{2,OP}$ at a higher rate of p as opposed to pq during Phase 4.

More precisely, for any $\epsilon, \delta > 0$, Tx_1 using random coding scheme of [27], encodes all the bits in $Q_{1,OP}$ and Q_{1,C_1} at rate $pq - \delta$ and communicates them. On the other hand, Tx_2 using random coding, encodes all the bits in $Q_{2\rightarrow 2|1}$ and all bits in $Q_{2,OP}$ at rate $p - \delta$. Then, Tx_2 communicates the XOR of its encoded bits.

Since Rx_1 already has access to the bits in $Q_{2\rightarrow 2|1}$ and $Q_{1,OP}$, it can remove the corresponding parts of the transmitted signals. Then for sufficiently large block length, Rx_1 can decode the transmitted bits from Q_{1,C_1} and $Q_{2,OP}$ with decoding error probability less than or equal to ϵ.

Finally, since Rx_2 already has access to the bits in $Q_{2,OP}$, it can remove the corresponding part of the transmitted signal. Then for sufficiently large block length, Rx_2 can decode the transmitted bits from $Q_{1,OP}$ and Q_{1,C_1} with decoding error probability less than or equal to ϵ. After decoding and removing this part, Rx_2 can decode the encoded bits from $Q_{2\rightarrow 2|1}$ with decoding error probability less than or equal to ϵ.

Phase 5 goes on for

$$\left[\frac{pq}{(1-q^2)}\left(p - q^2\right)m + 2m^{2/3}\right] / (p - \delta) \tag{116}$$

time instants. If an error occurs in decoding any of the encoded signals in Phase 5, we consider it as error and halt the transmission strategy. It is straight forward to verify that at the end of Phase 5, if the transmission is not halted, all states are empty and all bits are successfully delivered.

The probability that the transmission strategy halts at any point can be bounded by the summation of error probabilities of types I-II and the probability that an error occurs in decoding the encoded bits. Using Chernoff-Hoeffding bound and the results of [27], we can show that the probability that the transmission strategy halts at any point approaches zero for $\epsilon, \delta \rightarrow 0$ and $m \rightarrow \infty$. Moreover, the total transmission requires

$$\frac{1}{p}m + 6m^{2/3} \tag{117}$$

time instants. Thus, Tx_1 achieves a rate of $pq(1+q)$ while Tx_2 achieves a rate of p.

This completes the achievability proof of Theorem 2.

Appendix E
Proof of Lemma 2

In this appendix, we provide the proof of Lemma 2. We first derive the outer-bound and then we describe the achievability. The outer-bound on R_i is the same as in Section B.

Suppose there are encoders and decoders at the transmitters and receivers respectively, such that each receiver can decode both messages with arbitrary small decoding error probability as the block length goes to infinity. We have

$$
\begin{aligned}
& n(R_1 + R_2 - \epsilon_n) \\
& \stackrel{(a)}{\leq} I(W_1, W_2; Y_1^n | G^n) \\
& = H(Y_1^n | G^n) - H(Y_1^n | W_1, W_2, G^n) \\
& \stackrel{(b)}{=} H(Y_1^n | G^n) \\
& \stackrel{(c)}{\leq} \sum_{t=1}^{n} H(Y_1[t] | G^n) \\
& \leq \left(1 - q^2\right) n,
\end{aligned} \tag{118}
$$

where $\epsilon_n \to 0$ as $n \to \infty$; and (a) follows from the fact that the messages and G^n are mutually independent, Fano's inequality, and the fact that Rx_1 should be able to decode both messages; (b) holds since the received signal Y_1^n is a deterministic function of W_1, W_2, and G^n; and (c) follows from the fact that conditioning reduces entropy. Dividing both sides by n and let $n \to \infty$, we get

$$R_1 + R_2 \leq 1 - q^2. \tag{119}$$

Below, we provide the achievability proof of Lemma 2. Let $\mathrm{W}_i \in \{1, 2, \ldots, 2^{nR_i}\}$ denote the message of user i.

In [27], it has been shown that for a binary erasure channel with success probability p, and for any $\epsilon, \delta > 0$, as long as the communication rate is less than or equal to $p - \delta$, we can have decoding error probability less than or equal to ϵ.

Codebook generation is as follows. Transmitter i creates 2^{nR_i} ($R_1 = p - \delta$ and $R_2 = pq - \delta$) independent codewords where each entry of the codewords is an i.i.d. Bernoulli 0.5 RV. For message index j, transmitter i will send the j^{th} codeword. Note that we can view the channel from Tx_2 to Rx_1 as a binary erasure channel with success probability pq (whenever $G_{11}[t] = 0$ and $G_{21}[t] = 1$, we get a clean observation of $X_2[t]$). Therefore, since $R_2 = pq - \delta$, Rx_1 can decode W_2 with arbitrary small decoding error probability as $n \to \infty$ and remove X_2^n from its received signal. After removing X_2^n, we can view the channel from Tx_1 to Rx_1 as a binary erasure channel with success probability p (whenever $G_{11}[t] = 1$, we get a clean observation of $X_1[t]$). Therefore, since $R_1 = p - \delta$, Rx_1 can decode W_1 with arbitrary small decoding error probability as $n \to \infty$. Similar argument holds for Rx_2. This completes the achievability proof of corner point

$$(R_1, R_2) = (p, pq). \tag{120}$$

Similarly, we can achieve corner point

$$(R_1, R_2) = (pq, p). \tag{121}$$

Therefore with time sharing, we can achieve the entire region as described in Lemma 2.

Appendix F
Achievability Proof of Theorem 3: Corner Point $\left(1 - q^2, 0\right)$

By symmetry, it suffices to describe the achievability strategy for corner ponit

$$(R_1, R_2) = \left(1 - q^2, 0\right), \tag{122}$$

when transmitters have delayed knowledge of channel state information and noiseless output feedback links are available from the receivers to the transmitters.

Our achievability strategy is carried on over $b + 1$ communication blocks each block with n time instants. Transmitter one communicates fresh data bits in the first b blocks and the final block is to help receiver one decode its corresponding bits. Transmitter and receiver two act as a relay to facilitate the communication between transmitter and receiver one. At the end, using our scheme we achieve rate tuple $\frac{b}{b+1}\left(1 - q^2, 0\right)$ as $n \to \infty$. Finally, letting $b \to \infty$, we achieve the desired corner point.

Let W_1^j be the message of Tx_1 in block j, $j = 1, 2, \ldots, b$. We assume $\mathrm{W}_1^j = a_1^j, a_2^j, \ldots, a_m^j$. We set

$$n = \left(1 - q^2\right)^{-1} m + m^{2/3}. \tag{123}$$

Achievability strategy for block 1: At the beginning of the communication block, we assume that the bits at Tx_1 are in queue (or state) $Q_{1\rightarrow 1}^1$. At each time instant t, Tx_1 sends out a bit from $Q_{1\rightarrow 1}^1$, and this bit will leave this queue if at least one of the outgoing links from Tx_1 was equal to 1 at the time of transmission. On the other hand, Tx_2 remains silent during the first communication block. If at the end of the communication block, queue $Q_{1\rightarrow 1}^1$ is not empty, we declare error type-I and halt the transmission.

At the end of first block, using output feedback links, transmitter two has access to the bits of Tx_1 communicated in the first block. More precisely, Tx_2 has access to the bits of Tx_1 communicated in Cases $11, 12, 14$, and 15 during the first communication block. Note that the bits communicated in these cases are available at Rx_2 and have to be provided to Rx_1. Transmitter two transfers these bits to $Q_{2\rightarrow 1|2}^1$.

Assuming that the transmission is not halted, let $N_{2\rightarrow 1|2}^1$ denote the number of bits in queue $Q_{2\rightarrow 1|2}^1$. The transmission strategy will be halted and an error type-II will occur, if $N_{2\rightarrow 1|2}^1 > \mathbb{E}[N_{2\rightarrow 1|2}^1] + pqm^{\frac{2}{3}}$. From basic probability, we know that

$$\mathbb{E}[N_{2\rightarrow 1|2}^1] = \frac{\sum_{j=11,12,14,15} \Pr\left(\text{Case j}\right)}{1 - \sum_{i=9,10,13,16} \Pr\left(\text{Case i}\right)} m = (1 - q^2)^{-1} pqm. \tag{124}$$

Using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II and decreases exponentially with m.

Achievability strategy for block j, $j = 2, 3, \ldots, b$: In the communication block j, Tx_2 treats the bits in $Q_{2\rightarrow 1|2}^{j-1}$ as its message and it uses a random code of rate $pq - \delta$ to transmit them. Note that the channel from Tx_2 to Rx_1 can be modeled as a point-to-point erasure channel (any time $G_{11}[t] = 1$ or $G_{21}[t] = 0$, we consider an erasure has taken place). Hence from [27], we know that for any $\epsilon, \delta > 0$ and sufficiently large block length, a rate of $pq - \delta$ is achievable from Tx_2 to Rx_1 with decoding error probability less than or equal to ϵ. Note that at rate $pq - \delta$, both receivers will be able to decode and hence remove the transmitted signal of Tx_2 at the end of communication block. If an error occurs in decoding the transmitted signal of Tx_2, we consider it as error and halt the transmission strategy.

On the other hand, the transmission strategy for Tx_1 is the same as block 1 for the first b blocks (all but the last block). At the end of communication block j, using output feedback links, transmitter two has access to the bits of Tx_1 communicated in Cases $11, 12, 14$, and 15 during the communication block j. Transmitter two transfers these bits to $Q_{2\rightarrow 1|2}^j$. If at the end of the communication block, queue $Q_{1\rightarrow 1}^j$ is not empty, we declare error type-I and halt the transmission.

Assuming that the transmission is not halted, let $N_{2\rightarrow 1|2}^j$ denote the number of bits in queue $Q_{2\rightarrow 1|2}^j$. The transmission strategy will be halted and an error type-II will occur, if $N_{2\rightarrow 1|2}^j > \mathbb{E}[N_{2\rightarrow 1|2}^j] + pqm^{\frac{2}{3}}$. From basic probability, we know that

$$\mathbb{E}[N_{2\rightarrow 1|2}^j] = \frac{\sum_{j=11,12,14,15} \Pr\left(\text{Case j}\right)}{1 - \sum_{i=9,10,13,16} \Pr\left(\text{Case i}\right)} m = (1 - q^2)^{-1} pqm. \tag{125}$$

Using Chernoff-Hoeffding bound, we can show that the probability of errors of types I and II and decreases exponentially with m.

Achievability strategy for block $b+1$: Finally in block $b+1$, no new data bit is transmitted and Tx_2 only communicates the bits of Tx_1 communicated in the previous block in Cases $11, 12, 14$, and 15 as described before.

We can show that the probability that the transmission strategy halts at any point approaches zero as $m \rightarrow \infty$.

Decoding: At the end of block $j+1$, Rx_1 decodes the transmitted message of Tx_2 in block $j+1$ and removes it from the received signal. Together with the bits it has obtained during block j, it can decode message W_1^j. Using similar idea, Rx_2 uses backward decoding to cancel out interfernce in the previous blocks to decode all messages.

This completes the achievability proof for corner ponit

$$(R_1, R_2) = \left(1 - q^2, 0\right).$$

REFERENCES

[1] A. Vahid, M. Maddah-Ali, and A. Avestimehr, "Interference channel with binary fading: Effect of delayed network state information," in *49th Annual Allerton Conference on Communication, Control, and Computing*, pp. 894–901, 2011.

[2] A. Vahid, M. Maddah-Ali, and A. Avestimehr, "Binary fading interference channel with delayed feedback," in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, pp. 1882–1886, 2012.

[3] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, Sept. 1956.

[4] M. A. Maddah-Ali and D. N. T. Tse, "Completely stale transmitter channel state information is still very useful," in *Forty-Eighth Annual Allerton Conference on Communication, Control, and Computing*, Sept. 2010.

[5] A. Ghasemi, A. S. Motahari, and A. K. . Khandani, "On the degrees of freedom of X channel with delayed CSIT," in *2011 IEEE International Symposium on Information Theory Proceedings*, (Saint-Petersburg, Russia), pp. 909–912, July 2011.

[6] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region of the two-user MIMO broadcast channel with delayed CSI," Dec. 2010. arxiv.org/abs/1101.0306.

[7] H. Maleki, S. Jafar, and S. Shamai, "Retrospective interference alignment," *Arxiv preprint arXiv:1009.3593*, 2010.

[8] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "On the degrees of freedom of K-user SISO interference and X channels with delayed CSIT," *Submitted to IEEE Transactions on Information Theory, Available online at: Arxiv preprint arXiv:1109.4314*, 2011.

[9] R. Tandon, S. Mohajer, V. Poor, and S. Shamai, "Degrees of freedom region of the mimo interference channel with output feedback and delayed csit," *accepted for publication in IEEE Transactions on Information Theory*, 2012.

[10] C. Vaze and M. Varanasi, "The degrees of freedom region of the mimo interference channel with shannon feedback," *arXiv preprint arXiv:1109.5779*, 2011.

[11] C. Vaze and M. Varanasi, "The degrees of freedom region and interference alignment for the mimo interference channel with delayed csit," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4396–4417, 2012.

[12] A. S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A detreministic approach," *IEEE Transactions on Information Theory*, vol. 57, Apr. 2011.

[13] G. Bresler and D. Tse, "The two-user gaussian interference channel: a deterministic view," *European Transactions on Telecommunications*, vol. 19, no. 4, pp. 333–354, 2008.

[14] G. Bresler, A. Parekh, and D. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4566–4592, 2010.

[15] C. Suh and D. Tse, "Feedback capacity of the gaussian interference channel to within 2 bits," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2667–2685, 2011.

[16] A. Vahid, C. Suh, and A. S. Avestimehr, "Interference channels with rate-limited feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2788–2812, 2012.

[17] A. Avestimehr, A. Sezgin, and D. Tse, "Capacity of the two-way relay channel within a constant gap," *European Transactions on Telecommunications*, vol. 21, no. 4, pp. 363–374, 2010.

[18] A. Sezgin, A. Avestimehr, M. Khajehnejad, and B. Hassibi, "Divide-and-conquer: Approaching the capacity of the two-pair bidirectional gaussian relay network," *IEEE Transactions on Information Theory*, vol. 58, no. 4, pp. 2434–2454, 2012.

[19] N. T. Gaarder and J. K. Wolf, "The capacity region of a multiple-access discrete memoryless channenl can increase with feedback," *IEEE Transactions on Information Theory*, vol. 21, pp. 100–102, Jan. 1975.

[20] L. Ozarow, "The capacity of the white gaussian multiple access channel with feedback," *IEEE Transactions on Information Theory*, vol. 30, no. 4, pp. 623–629, 1984.

[21] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, 1963.

[22] J. Phillips, "Chernoff-hoeffding inequality and applications," *arXiv preprint arXiv:1209.6396*, 2012.

[23] B. Nazer, M. Gastpar, S. A. Jafar, and S. Vishwanath, "Ergodic interference alignment," *Arxiv.org*, 2011. avilable at arxiv.org/abs/0901.4379.

[24] T. Gou, S. A. Jafar, S. W. Jeon, and S. Y. Chung, "Aligned interference neutralization and the degrees of freedom of the $2 \times 2 \times 2$ interference channel," in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, pp. 2751–2755, 2011.

[25] I. Shomorony and A. S. Avestimehr, "Two-unicast wireless networks: Characterizing the degrees-of-freedom," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 353–383, 2013.

[26] I. Shomorony and A. S. Avestimehr, "Degrees of freedom of two-hop wireless networks: "everyone gets the entire cake"," in *50th Annual Allerton Conference on Communication, Control, and Computing*, 2012.

[27] P. Elias, "The noisy channel coding theorem for erasure channels," *The American Mathematical Monthly*, pp. 853–862, 1974.